EDITED BY

BERND
HEINE

HEIKO
NARROG

The Oxford Handbook *of* LINGUISTIC ANALYSIS

THE OXFORD HANDBOOK OF

LINGUISTIC ANALYSIS

OXFORD HANDBOOKS IN LINGUISTICS

Published

Comparative Syntax
Edited by Gugliemo Cinque and Richard S. Kayne

Cognitive Linguistics
Edited by Dirk Geeraerts

Linguistic Analysis
Edited by Bernd Heine and Heiko Narrog

Compounds
Edited by Rochelle Lieber and Pavol Štekauer

Applied Linguistics
Edited by Robert B. Kaplan

Case
Edited by Andrej Malchukov and Andrew Spencer

Japanese Linguistics
Edited by Shigeru Miyagawa and Mamoru Saito

Computational Linguistics
Edited by Ruslan Mitkov

Linguistic Interfaces
Edited by Gillian Ramchand and Charles Reiss
[Published in Association with Oxford Studies in Theoretical Linguistics]

THE OXFORD HANDBOOK OF

LINGUISTIC ANALYSIS

Edited by

BERND HEINE

and

HEIKO NARROG

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2010

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire

ISBN 978–0–19–954400–4

1 3 5 7 9 10 8 6 4 2

Dedicated to John Davey

This page intentionally left blank

Contents

This page intentionally left blank

List of Abbreviations

1, 2, 3	first, second, third person
2SO	second person singular object agreement
2SS	second person singular subject agreement
3SO	third person singular object agreement
3SS	third person singular subject agreement
A	agent; adjective
ABIL	ability
ABL	ablative
ABS	absolute, absolutive
ACC, Acc	accusative, operator of acceptability
ADJ	adjective; adjunct
ADV, Adv	adverb
Agr	agreement
AgrS	subject–verb agreement
AmE	American English
ANIM	animate
ANTIPASS	antipassive
AP	adverbial phrase
ASL	American Sign Language
ASP	aspect
AUH	actor–undergoer hierarchy
AUX	auxiliary
AVM	attribute-value matrix
BA	bare argument
BAE	bare argument ellipsis
BEN	beneficiary
BNC	British National Corpus
BrE	British English

BSL	British Sign Language
C	consonant; complementizer
C1, C2, etc.	noun class 1, 2, etc.
CAT	syntactic category
CCM	Constituent-Context Model
CD	cognitive default
CG	Cognitive Grammar
CL	class, Cognitive Linguistics
CN	count noun
CNJ	conjunction
COMP	complement
COMPL	completive
Con	component that defines the set of universal violable constraints
COND	conditional
Conj	conjunction
CONN	connective
CONTR	contrastive
COP	copula
COREF	coreference
CORR	correlative
CP	complementizer phrase
CPI	effortful processing
CS	conceptual structure
CSLI	Center for the Study of Language and Information
D	determiner; dependency
DAT, Dat	dative
DECL	declarative
DEF	definite
DEM, Dem	demonstrative
DET	determiner
DG	dependency grammar
DI	degree of intention
DO	direct object
DOP	Data-Oriented Parsing

DS	Default Semantics
DU	Derivational Uniformity
DUP	duplicative
ECG	Embodied Construction Grammar
epf	epistemic possibility future
ERG	ergative
Eval	component that selects an optimal output from the set of alternative candidates
F	feminine; Finite
f	form pole of a construction
FACT	factual mood
FAM	familiar
FDEF	type of faithfulness constraint
FDG	Functional Discourse Grammar
FDR	type of faithfulness constraint
FE	frame element
FEM	feminine
FG	Functional Grammar
FGT	Formal Generative Typology
fif	fully inflected form
FL	faculty of language
FN	FrameNet
FocP	Focus projection
fp	futurative progressive
FPL	type of faithfulness constraint
fr	frame
*FUNCTN	a core markedness constraint
FUT	future
GAPP	Golden Age of Pure Pragmatics
GB	Government-Binding Theory
GCat	morphological constituents
GCI	generalized conversational implicature
GEN, Gen	genitive
Gen	component where output candidate parses are generated

GF	grammatical function
GPSG	Generalized Phrase Structure Grammar
H	high tone
HAB	habitual
HPSG	Head-Driven Phrase Structure Grammar
I	inflection
IF	illocutionary force
IFG	Incremental Functional Grammar
IL	interpersonal level
IMP	imperative
IMPF	imperfective aspect
IND	indicative mood
INDEF	indefinite
INDIC	indicative mood
INF	infinitive
INFER	inferential
Infl	inflection
INSTR	instrument(al)
IO	indirect object
IP	inflectional phrase
IRR	irrealis
IS	information structure, inferential system
isa	basic relation of classification
IU	Interface Uniformity
JSL	Japanese Sign Language
K	case
L	low tone; loser
LAD	language acquisition device
LDP	left-detached position
LF	logical form
LFG	Lexical Functional Grammar
LIS	Italian Sign Language
LOC	locative
LS	logical structure

LSC	Catalan Sign Language
LSF	French Sign Language
LU	lexical unit
M	masculine; semantic molecule
MASC	masculine
MAX	constraint requiring every segment in the input to have a correspondent in the output
MD	Multi-Dimensional analysis
MDP	Minimal Distance Principle
MGG	Mainstream Generative Grammar
MP	Minimalist Program
N	noun; neuter
NAND	"and not"
NEG	negation
NMR	non-macrorole
NOGEM	constraint against geminates
NOM, Nom	nominative
NONPST	nonpast
NP	noun phrase
NSM	Natural Semantic Metalanguage
NTL	Neural Theory of Language
NUC	nucleus
NUM, Num	numeral
O, OBJ	object
OBL	oblique
orth	orthography
OT	Optimality Theory
OT-LFG	Optimality-Theoretic Lexical Functional Grammar
P	patient; Predicator
PA	Parallel Architecture
P&P	Principles-and-Parameters
ParGram	Parallel Grammar
ParSem	Parallel Semantics
PART	participle

PASS	passive
PCat	levels in the Prosodic Hierarchy
PCFG	Probabilistic Context-Free Grammar
PCI	particularized conversational implicature
PERF	perfect
pers	person
PF	perfective
PI	primary intention
PL, pl	plural
pm	primary meaning
POSS	possessive
PP	prepositional phrase
PrCS	pre-core slot
PRED	predicate
PRES	present
PRO	pronoun
PROG	progressive
PRON, Pron	pronoun
PROX	proximative aspect
PRS	present
PRT	particle
PSA	privileged syntactic argument
PST	past
PTCP	participle
PUNC	punctual aspect
PW	phonological word
R	form–government relation
REAL	realis
REC	reciprocal
REF	reflexive
REF-REC	reflexive-reciprocal category
REL	relative clause marker
REM	remote
REP	reportative

RG	Relational Grammar
RL	representational level
RP	reference phrase
RRG	Role and Reference Grammar
RT	relevance theory
S	sentence; subject
SAI	subject–auxiliary inversion
S&W	Sperber and Wilson
SBCG	Sign-Based Construction Grammar
SBJV	subjunctive mood
SC	society and culture
SCWD	automatic utilization of knowledge of culture and society
SD	situation of discourse
SEM	meaning of a sign
SFG	Systemic Functional Grammar
SFL	Systemic Functional Linguistics
SG, sg	singular
sm	secondary meaning
SOV	subject–object–verb
SP	subject pronoun
SPE	*The Sound Pattern of English*
SPEC	specifier
SS	Simpler Syntax
SSH	Simpler Syntax Hypothesis
STAT	stative aspect
SU	Structural Uniformity
SUBJ	subject
SUBS	subsequent tense
SVO	subject–verb–object
T	tense; word-token
TAG	Tree-Adjoining Grammar
TAM	tense–aspect–modality
TL	trajectory-landmark
TMA	tense–modality–aspect

TNS	tense
TOP	topic
TR	transitive
TSG	Tree-Substitution Grammar
UG	Universal Grammar
UTAH	Uniformity of Theta Assignment Hypothesis
V	verb; vowel
VAL	valence
VIS	visual
VOC	Verb–Object Constraint
VP	verb phrase
VSO	verb–subject–object
VT	valency theory
W	winner
WALS	*World Atlas of Language Structures*
W&S	Wilson and Sperber
WG	Word Grammar
WH	question word
WK	world knowledge
WS	word meaning and sentence structure
XADJ	open predicate adjunct
XCOMP	open predicate complement
XLE	Xerox Linguistic Environment
XOR	a set-theoretic operator inspired by exclusive disjunction
XP	predicate phrase (VP, NP, AP, or PP)

About the Authors

Vilmos Ágel is Chair of German Synchronic Linguistics at Kassel University. He received his doctorate (1988) and professorial degree (1997) from the ELTE Budapest. The bursaries he received include the following: 1987–8, from the German Academic Exchange Service, 1991–3 and 1998, Alexander von Humboldt-Stiftung, and 2000–4, Szécheny research bursary for professors. In 2003, he won the Friedrich Wilhelm Bessel research prize of the Alexander von Humboldt-Stiftung. Since 2004 he has been co-editor of the *Zeitschrift für Germanistische Linguistik*. His main research interests are valency theory, dependency grammar, the link between contemporary grammar and the history of language, the grammar of New High German (1650–2000), and the relationship between orality/literacy and grammatical structure.

Ash Asudeh teaches linguistics and cognitive science at Carleton University in Ottawa, Canada. He received his Ph.D. from Stanford University in 2004. His primary research interests are syntax, semantics, the intersection of logic and linguistics, and grammatical architecture.

Collin Baker received a Ph.D. in Linguistics from University of California at Berkeley in 1999, and since 2000 has been manager of the FrameNet Project at the International Computer Science Institute. Recent publications include *The Structure of the FrameNet Database* (2003) (with Charles Fillmore and Beau Cronin) and *FrameNet and Representing the Link between Semantic and Syntactic Relations* (2004) (with Fillmore and Josef Ruppenhofer), along with a number of conference papers and tutorials at Language Resources and Evaluation Conference (LREC), Association for Computational Linguistics (ACL), and Linguistic Society of America (LSA). His current research focuses on efficiently extending the FrameNet lexicon and on annotating large corpora with semantic frames.

Mark Baker is a professor in the Department of Linguistics and the Center for Cognitive Science at Rutgers University. He received his Ph.D. in linguistics from MIT in 1985, and has also taught at McGill University. He is the author of numerous technical articles and four research monographs: *Incorporation* (1988), *The Polysynthesis Parameter* (1996), *The Lexical Categories* (2003), and *The Syntax of Agreement and Concord* (2008). He has also written one book for a more general audience, *The Atoms of Language* (2001). His primary research interest is the syntax and morphology of less-studied languages, especially Native American languages and African languages.

Clay Beckner is a Ph.D. candidate in Linguistics at the University of New Mexico. In his research, he studies usage-based language change with a variety of methodologies, including corpus-based statistical methods, psycholinguistics, and computational modeling. He is working on a dissertation that addresses the formation of multiword units in language, incorporating experimental data and exemplar models of usage and change.

Douglas Biber is Regents' Professor of English (Applied Linguistics) at Northern Arizona University. His research efforts have focused on corpus linguistics, English grammar, and register variation (in English and cross-linguistic; synchronic and diachronic). His publications include books on register variation and corpus linguistics published by Cambridge University Press (1988, 1995, 1998, to appear), the co-authored *Longman Grammar of Spoken and Written English* (1999), and more recent studies of language use in university settings and discourse structure investigated from a corpus perspective (both published by Benjamins: 2006 and 2007).

Rens Bod obtained his Ph.D. in Computational linguistics at the University of Amsterdam. He currently holds a chair in Artificial Intelligence and is affiliated with the University of St Andrews and at the University of Amsterdam. His main interests are in computational models of language acquisition and statistical natural language processing, but he has also worked in computational musicology, reasoning, and in the history of science and humanities. Rens published over 90 scholarly articles and is the author of *Beyond Grammar: An Experience-Based Theory of Language* (Stanford University, 1998). He also co-edited a number of handbooks, including *Probabilistic Linguistics* (MIT Press, 2003) and *Data-Oriented Parsing* (CSLI Publications/Chicago, 2003).

Cedric Boeckx is Research Professor at the ICREA (Institució Catalana de Recerca i Estudis Avançats) and a member of the CLT (Centre de Lingüística Teòrica) at the Universitat Autònoma de Barcelona. He received his Ph.D. in linguistics in 2001 from the University of Connecticut, Storrs. He has held visiting positions at the Universities of Illinois and Maryland, and was a fellow of Netherlands Organization for Scientific Research (NWO) at Leiden University. His research interests are in theoretical syntax, comparative grammar, and architectural questions of language, including its origins and its development in children and its neurobiological basis. He is the author of, among others, *Islands and Chains* (Benjamins 2003), *Linguistic Minimalism* (Oxford University Press, 2006), *Understanding Minimalist Syntax* (Blackwell, 2008), and *Bare Syntax* (Oxford University Press, 2008). He has published numerous articles in journals such as *Linguistic Inquiry* and *Natural Language and Linguistic Theory*.

Geert Booij was born in 1947 in Hoogeveen, The Netherlands. In 1971, he received an M.A. (cum laude) in Dutch linguistics, with minors in general linguistics and philosophy from the University of Groningen. From 1971 until 1981 he was an assistant/associate professor in the Department of Dutch of the University of Amsterdam, where he obtained his Ph.D. degree in 1977 with the dissertation *Dutch Morphology. A Study of Word Formation in Generative Grammar* (Foris Publications, Dordrecht). From 1981 until 2005, Geert Booij was a full professor of General Linguistics at the Vrije Universiteit Amsterdam. As of September 2005, he is Professor of Linguistics at the University of Leiden. Geert Booij is, with Jaap van Marle, the founder and editor of the book series *Yearbook of Morphology* which is now, as of 2006, the journal *Morphology*. He is the author of *The Phonology of Dutch* (1995), *The Morphology of Dutch* (2002), and of *The Grammar of Words* (2005), all published by Oxford University Press, and of a number of linguistic articles in a wide range of Dutch and international journals.

John Bryant is a researcher at the International Computer Science Institute. He received his Ph.D. in computer science at the University of California at Berkeley in 2008. His research investigates precise, cognitively motivated implementations of language interpretation using constructions.

Joan L. Bybee (Ph.D. linguistics, University of California at Los Angeles, 1973) was on the faculty at the State University of New York at Buffalo from 1973 until 1989 and is now Distinguished Professor Emerita of the Department of Linguistics at the University of New Mexico. Bybee's research interests include theoretical issues in phonology and morphology, language universals and linguistic change. Her books include *Morphology* (1985), *The Evolution of Grammar* (1994) (with Revere Perkins and William Pagliuca), *Phonology and Language Use* (2001) and *Frequency of Use and the Organization of Language* (2006). In 2004 she served as the President of the Linguistic Society of America.

Alice Caffarel is Senior Lecturer in the School of Languages and Cultures at the University of Sydney, where she teaches French, French linguistics and stylistics. She is the author of a Systemic Functional Grammar of French and the co-editor of a volume on language typology from a systemic functional perspective.

Guglielmo Cinque is Professor of Linguistics at the University of Venice, co-editor of *The Oxford Handbook of Comparative Syntax* (Oxford University Press, 2005) (with Richard Kayne), and author of various volumes on syntax, including *Adverbs and Functional Heads* (Oxford University Press, 1999) and *Restructuring and Functional Heads* (Oxford University Press, 2006). With Luigi Rizzi he has promoted the subseries of volumes on *The Cartography of Syntactic Structures* of the Oxford Studies in Comparative Syntax series edited by R. Kayne.

Peter Culicover received his Ph.D. in Linguistics from MIT in 1971. He is currently Humanities Distinguished Professor in Linguistics at the Ohio State University. He was Chair of the Department of Linguistics for eight years and was the founding Director of the Center for Cognitive Science at OSU and served in that capacity for thirteen years. He has published twelve books and over sixty articles and chapters. In 2006, he was awarded the Humboldt Research Award by the Alexander von Humboldt Foundation and is a recent winner of the Distinguished Scholar Award of the Ohio State University.

Henriëtte de Swart obtained her Ph.D. from the University of Groningen in 1991. She has been full Professor of French Linguistics and Semantics at Utrecht University since 1997. The main emphasis of her work is on cross-linguistic semantics. Her collaboration with Donka Farkas led to a book on noun incorporation, and several articles on genericity, plurality, and indefiniteness. In recent years, she has explored Optimality Theory to develop typologies of negation and article use. With Joost Zwarts, she is currently involved in a research project on "weakly referential noun phrases".

Ellen K. Dodge is a Ph.D. candidate in Linguistics at the University of California, Berkeley. Her main area of research is in cognitive linguistics, and includes work on image schemas, meaning representation, construction grammar, and compositionality.

Jerome A. Feldman is a Professor of Computer Science and Cognitive Science at the University of California and a research scientist at the International Computer Science Institute in Berkeley. He received the Ph.D. from Carnegie-Mellon in 1964 and previously taught at Stanford and Rochester.

Charles F. Fillmore is Emeritus Professor at the University of California at Berkeley. His early work included contributions to generative grammar (cyclic rule application), the grammar and semantics of English verbs, case grammar, deixis, frame semantics, and construction grammar. After retirement he has been associated with FrameNet, a computational lexicography project at the International Computer Science Institute, supported mainly by the US National Science Foundation. He served as President of the Linguistic Society of America in 1991.

Klaus Fischer started his academic career in Britain as *DAAD-Lektor* (Deutscher Akademischer Austauschdienst) at the University of Wales, Lampeter (1985–88), after having studied German and Philosophy in Munich and Bonn. He has worked at London Metropolitan University or one of its predecessor institutions (City of London Polytechnic, London Guildhall University) since 1989, first as Lecturer in German, heading the German section since 1999, and from 2002 as Reader in German. In 1995, he was awarded a Ph.D. at the University of Wales. His main research interests are the foundation and application of valency theory, the comparison of English and German verb complementation, and more generally the typology of English and German.

T. Givón is Professor Emeritus of Linguistics and Cognitive Science and co-founder of the Institute of Cognitive and Decision Sciences at the University of Oregon. From 1969 to 1981 he taught at the University of California at Los Angeles, and from 1981 to 2004 at the University of Oregon. His many books include *On Understanding Grammar* (Academic Press, 1979); *Ute Dictionary* (Ute Press, 1979), *Ute Reference Grammar* (Ute Press, 1980), *Ute Traditional Narratives* (Ute Press, 1985), *Syntax: A Functional-Typological Introduction* (Benjamins, 1984, 1990), *Mind, Code and Context* (Erlbaum, 1989), *English Grammar* (Benjamins, 1993), *Functionalism and Grammar* (Benjamins, 1995), and *Syntax: An Introduction* (Benjamins, 2001). His most recent books are *Bio-Linguistics: The Santa Barbara Lectures* (Benjamins, 2002) and *Context as Other Minds: The Pragmatics of Sociality, Cognition and Communication* (Benjamins, 2005).

Cliff Goddard's research interests lie at the intersection of language, meaning and culture. He is one of the leading proponents of the NSM (Natural Semantic Metalanguage) approach to linguistic description and a close collaborator with the originator of the approach, Anna Wierzbicka. He has published extensively on topics in semantic theory, cross-linguistic semantics, cross-cultural pragmatics, and on language description, especially in relation to English, Malay, and Yankunytjatjara (Central Australia). Recently he edited the collective volumes *Ethnopragmatics: Understanding Discourse in Cultural Context* (Mouton de Gruyter, 2006) and *Cross-Linguistic Semantics* (Benjamins, 2008). With Anna Wierzbicka, he is co-editor of the two-volume collection *Meaning and Universal Grammar–Theory and Empirical Findings* (Benjamins, 2002). Goddard has published two textbooks: *Semantic Analysis* (OUP, 1998) and *The Languages of East and Southeast Asia* (OUP, 2005). He is a full Professor of Linguistics at the University of New England, Australia, and a Fellow of the Australian Academy of Humanities.

Maria Gouskova is an Assistant Professor in the Department of Linguistics at New York University. She works in phonology and its interfaces with phonetics and morphology in the framework of Optimality Theory. She is interested in the relation of phonological constraints to each other and to linguistic primitives. She has published in the journals *Linguistic Inquiry* and *Phonology* on a variety of topics in phonological theory ranging from the role of sonority in syllable structure to templates in reduplication.

Martin Haspelmath is a senior staff member at the Max-Planck-Institut für evolutionäre Anthropologie in Leipzig. He received degrees at the Universität zu Köln, University at Buffalo, and the Freie Universität Berlin. He taught at the Freie Universität (FU) Berlin, the Universität Bamberg, the Università di Pavia, the Universität Leipzig, and at summer schools in Albuquerque, Mainz, Düsseldorf, Cagliari, Berkeley, and at MIT. His research interests are primarily in the area of broadly comparative and diachronic morphosyntax (*Indefinite Pronouns*, 1997; *From Space* *to Time*, 1997; *Understanding Morphology*, 2002). He is one of the editors of Oxford University Press's *World Atlas of Language Structures* (2005).

Bernd Heine is Emeritus Professor at the Institute of African Studies (Institut für Afrikanistik), University of Cologne. His 33 books include *Possession: Cognitive Sources, Forces, and Grammaticalization* (Cambridge University Press, 1997); *Auxiliaries: Cognitive Forces and Grammaticalization* (Oxford University Press, 1993); *Cognitive Foundations of Grammar* (Oxford University Press, 1997); with Derek Nurse, *African Languages: An Introduction* (Cambridge University Press, 2000); *A Linguistic Geography of Africa* (Cambridge University Press, 2008); with Tania Kuteva, *World Lexicon of Grammaticalization* (Cambridge University Press, 2002); *Language Contact and Grammatical Change* (Cambridge University Press, 2005); *The Changing Languages of Europe* (Oxford University Press, 2006).

Kees Hengeveld is Professor of Theoretical Linguistics at the University of Amsterdam. Before that, he taught Spanish Linguistics at that same university. His research focuses on Functional Discourse Grammar and linguistic typology, and often on the combination of the two. With Lachlan Mackenzie he published *Functional Discourse Grammar* (Oxford University Press, 2008), a full account of the model presented in outline form in the present book. Before that, he posthumously edited Simon C. Dik's two-volume *The Theory of Functional Grammar* (Mouton de Gruyter, 1997), and published *Non-Verbal Predication: Theory, Typology, Diachrony* (Mouton de Gruyter, 1992), as well as numerous articles and edited volumes.

Yan Huang is Professor of Linguistics at the University of Auckland. He has previously taught linguistics at the University of Cambridge, University of Oxford, and University of Reading, where he was Professor of Theoretical Linguistics. His main research interests are in pragmatics, semantics and syntax, especially the pragmatics–semantics interface and the pragmatics–syntax interface. His published work includes *The Syntax and Pragmatics of Anaphora* (Cambridge University Press, 1994, re-issued 2007), *Anaphora: A Cross-Linguistic Study* (Oxford University Press, 2000), and *Pragmatics* (Oxford University Press, 2007). He has also published a number of articles and reviews in leading international journals of linguistics.

Richard ("Dick") Hudson was born in 1939 and educated at Loughborough Grammar School, Corpus Christi College Cambridge and the School of Oriental and African Studies (SOAS), London. Since his 1964 SOAS Ph.D., which dealt with the grammar of the Cushitic language Beja, he spent all his salaried research life working on English at University College London, with occasional forays into other languages. Another strand of his linguistics, due to early contacts with Michael Halliday, was (and is) an attempt to improve the bridge between academic linguistics and school-level language education. He was elected Fellow of the British Academy in 1993 and retired from UCL in 2004.

Ray Jackendoff is Seth Merrin Professor of Philosophy and Co-Director of the Center for Cognitive Studies at Tufts University. He is also Professor Emeritus of Linguistics at Brandeis University and a member of the External Faculty of the Santa Fe Institute. He is past President of the Linguistic Society of America and the 2003 recipient of the Jean Nicod Prize in cognitive philosophy. His most recent books are *Foundations of Language*, *Simpler Syntax* (2005) (with Peter Culicover), and *Language, Consciousness, Culture* (2007).

Kasia M. Jaszczolt is Reader in Linguistics and Philosophy of Language at the University of Cambridge and Fellow of Newnham College, Cambridge. She obtained her D.Phil. from the University of Oxford (1992) for a thesis on the semantics of propositional attitude reports. She has published widely on the topics of semantic ambiguity and underdetermination, definite descriptions, belief reports, and the semantics of time, principally developing and applying the theory of Default Semantics to various types of constructions. Her books include *Default Semantics* (Oxford University Press, 2005), *Semantics and Pragmatics: Meaning in Language and Discourse* (Longman, 2002), *Discourse, Beliefs and Intentions* (Elsevier Science, 1999) and *Representing Time: An Essay on Temporality as Modality* (forthcoming from Oxford University Press in 2008). She is also member of various editorial boards including *Journal of Pragmatics* and *Studies of Pragmatics*, a committee member of the Linguistics Association of Great Britain, and in from 1996 to 2008 was managing editor of a book series *Current Research in the Semantics/Pragmatics Interface* (*CRiSPI*) (Elsevier Science).

Ronald W. Langacker is Professor Emeritus and Research Professor at the University of California, San Diego. Early in his career, major areas of research included generative syntax and the comparative-historical grammar of the Uto-Aztecan family of Native American languages. Over the last three decades, his main concern has been to develop and progressively articulate the theory of Cognitive Grammar, a radical alternative to generative theory. He has published numerous books and articles dealing with a broad array of issues in cognitive linguistics.

J. Lachlan Mackenzie is Honorary Professor of Functional Linguistics at the Free University of Amsterdam, the Netherlands, and works at Instituto de Linguística Teórica e Computacional (Lisbon, Portugal) as a Consultant in Languages and Linguistics. His professional interests are mainly in the analysis of Western European languages from the perspective of Functional Discourse Grammar (FDG). Among recent books co-edited in this area are *A New Architecture for Functional Grammar* (Mouton de Gruyter, 2004), *Crucial Readings in Functional Grammar* (Mouton de Gruyter, 2005), and *Studies in Functional Discourse Grammar* (Lang, 2005). He is co-author, with Kees Hengeveld, of *Functional Discourse Grammar* (Oxford University Press, 2008).

Laura A. Michaelis is an Associate Professor of Linguistics and a faculty fellow in the Institute of Cognitive Science at the University of Colorado at Boulder. She is the author of *Aspectual Grammar and Past-Time Reference* (1998) and *Beyond Alternations: A Constructional Model of the German Applicative Pattern* (2001) (with Josef Ruppenhofer). She is the co-editor, with Elaine Francis, of *Mismatch: Form-Function Incongruity and the Architecture of Grammar* (2003). She is currently collaborating on a Construction Grammar monograph with Charles Fillmore, Paul Kay, and Ivan Sag. Her work has appeared in the journals *Language*, *Journal of Semantics*, *Journal of Pragmatics, Cognitive Linguistics, Journal of Linguistics, Linguistics and Philosophy*, and *Studies in Language.*

Glyn Morrill has held research positions at the Centre for Cognitive Science, University of Edinburgh, Netherlands Organization for Scientific Research (NWO) Universiteit Utrecht and Centrum voor Wiskunde en Informatica (CWI) Amsterdam, and Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya. Since 1993 he has been a lecturer in Computer Science at this latter. In 1995, he co-founded with Dick Oehrle the annual Formal Grammar conference. He researches on logic in grammar and processing, focusing on type logical categorial grammar. He is author of *Type Logical Grammar: Categorial Logic of Signs* (Kluwer, 1994) and *Lògica de primer ordre* (Edicions Universitat Politecnica de Catalunya, 2001).

Heiko Narrog is Associate Professor at the Graduate School of International Cultural Studies of Tohoku University. He holds two Ph.D.s in linguistics, in Germany and Japan, and his publications include *Japanische Verbflexive und flektierbare Suffixe* (Harrassowitz, 1999) as well as numerous articles in linguistic typology, semantics and language change, and Japanese linguistics. He is currently involved in a typological project on semantic maps and is preparing the publication of a book on modality and the hierarchy of functional categories.

William O'Grady is Professor of Linguistics at the University of Hawaii at Manoa. His recent work, exemplified by his 2005 book, *Syntactic Carpentry* (Erlbaum), adopts an emergentist perspective on syntax, arguing that an efficiency-driven processor can explain why language has the particular properties that it does and how those properties are acquired with such apparent ease by children. He is also well known for his work on Korean, on which he has written extensively.

Eric Pederson received his Ph.D. in linguistics from the University of California, Berkeley, in 1991 with a dissertation on the polysemy of reflexive and causative morphology. He then worked with the Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics in Nijmegen through 1997. His work has long focused on spatial language and cognition as well as the development of field methods relevant for linguistic relativity research. Since 1984, Pederson's field work has been mostly in rural South India. In Winter 1998, he joined the Linguistics

Department of the University of Oregon, where he is currently an Associate Professor and the Department Head.

Luigi Rizzi is Professor of Linguistics at the University of Siena. He has been on the faculty of several universities in Europe and the US, including MIT, the University of Geneva, the Ecole Normale Supérieure (Paris). His research focuses on syntactic theory and comparative syntax, with special reference to the theory of locality, the study of variation through parametric models, the cartography of syntactic structures, the acquisition of syntax.

Ida Toivonen teaches linguistics and cognitive science at Carleton University in Ottawa, Canada. She received her Ph.D. from Stanford University in 2001. Her primary area of research is syntax, but she also works on phonetics, morphosyntax, lexical semantics, and language description. She is particularly interested in Swedish and Inari Saami.

Robert D. Van Valin, Jr. received his Ph.D. in linguistics at the University of California, Berkeley. His research is focused on theoretical linguistics, especially syntactic theory and theories of the acquisition of syntax and the role of syntactic theory in models of sentence processing. He is the co-author of *Functional Syntax and Universal Grammar* (Cambridge University Press, 1984), the editor of *Advances in Role and Reference Grammar* (Benjamins, 1993), the co-author of *Syntax: Structure, Meaning and Function* (Cambridge University Press, 1997), and the author of *An Introduction to Syntax* (Cambridge University Press, 2001). His most recent book is *Exploring the Syntax–Semantics Interface* (Cambridge University Press, 2005). He is the general editor of the *Oxford Surveys in Syntax and Morphology* series (Oxford University Press).

Phyllis Perrin Wilcox is Professor in the Department of Linguistics at the University of New Mexico. She has presented on metaphor and metonymy in signed languages at national and international conferences. Dr. Wilcox was instrumental in establishing the baccalaureate degree in Signed Language Interpreting at UNM and coordinated the program for over fifteen years. She has served as chair of the Registry of Interpreters for the Deaf's National Review Board and now serves on the New Mexico Interpreting Licensure Board. Dr. Wilcox is the author of *Metaphor in American Sign Language* (Gallaudet University Press, 2000) and numerous articles on metaphorical cognition.

Sherman Wilcox is Professor of Linguistics, University of New Mexico. Dr. Wilcox's interests include understanding what signed languages can tell us about the origin of language, as well as the developmental process by which gesture becomes incorporated into the grammatical system of signed languages. Dr. Wilcox is the author of many articles on a wide range of topics focussing on the theoretical and applied study of signed languages. His books include *Gesture and the Nature of Language*

(Cambridge University Press, 1995) (with David Armstrong and William Stokoe), *Learning to See: Teaching American Sign Language as a Second Language* (Gallaudet, 1997) (with Phyllis Perrin Wilcox), and *The Gestural Origin of Language* (Oxford University Press, 2007) (with David Armstrong).

Francisco Yus teaches pragmatics at the University of Alicante, Spain. He has a Ph.D. in linguistics and has specialised in the application of pragmatics (especially relevance theory) to media discourses and conversational issues. He has applied pragmatics to alternative comics (*Conversational Cooperation in Alternative Comics*, 1995; *El discurso femenino en el cómic alternativo inglés*, 1998), proposed a verbal-visual model of communication in media discourses (*La interpretación y la imagen de masas*, 1997), and developed a pragmatic approach to Internet-mediated communication (*Ciberpragmática*, 2001). Latest research has to do with the application of relevance theory to misunderstandings and irony in conversation, as well as to the production and interpretation of humor.

Joost Zwarts defended his Ph.D. thesis at Utrecht University in 1992. He is an Assistant Professor in the Department of Linguistics at Utrecht University and works in the area of semantics, with a special focus on the domains of spatial prepositions, article use, and lexical semantics. He has also done fieldwork in Kenya, and published a phonology of Endo, a Southern Nilotic language. He is now involved, with Henriëtte de Swart, in a research project on "weakly referential noun phrases".

CHAPTER 1

INTRODUCTION

BERND HEINE
HEIKO NARROG

LIKE the other volumes of the *Oxford Handbook in Linguistics* series, the present volume aims at offering "an authoritative and state-of-the art survey of current thinking and knowledge in a particular field" to serve as a source of reference for scholars and graduate students. Its format, however, differs from that of most other volumes of the series. The volume does not really have the internal structure that one might expect a handbook to have. Rather than grouping the chapters according to a catalog of more general themes, the table of contents has the format of a "shallow" taxonomy, simply listing the chapter titles and contributors. The editors have given the question of how the various chapters should be arranged and presented in a volume of this kind quite some thought. In the end they decided to simply arrange the chapters in alphabetical order of the first key word figuring in the chapter title, for the following reason: Current linguistic analysis has turned into an extremely complex field and imposing a rigid classification of theoretical concepts and orientations has become increasingly difficult and controversial. The editors therefore came to the conclusion that it would be best to leave it to the reader to find his or her own way in comparing and relating the chapters to one another. We are aware that this is a procedure that is not really the one expected from a handbook-type treatment but we believe that it suggests itself on the basis of the nature of the volume.

There are also other ways in which the volume differs from handbook-type works. "Linguistic analysis" is a fairly general notion, being suggestive of a

We wish to express our gratitude to Christopher Collins, Fritz Newmeyer, and Frans Plank for valuable comments on an earlier version of this chapter.

comprehensive treatment of what contemporary linguistics has to offer for analyzing human languages. Yet, the reader looking for information on methodologies for analyzing, for example, phonology or morphology, two of the liveliest areas of inquiry in recent linguistics, may be disappointed, since these areas are not represented to the extent they should be in a volume of this nature. The editors decided to give priority to approaches that focus on morphosyntax, the interaction of morphology and syntax, which is widely believed to be the most central domain of grammar, allowing for the largest range of generalizations on language structure (cf. Haspelmath, this volume, who equates "grammatical theory" with "theory of morphosyntax"). It goes without saying that this is not the only possible procedure that could have been adopted.

A major objective that the editors had in mind when embarking on the handbook project was to give those scholars who were responsible for, or are prominently involved in, the development of a given approach, program, or theory a chance to describe and promulgate their work. Another objective was that, rather than offering a limited selection of mainstream lines of linguistics, we wanted to expose the reader to a broad range of theoretical discussions. To this end, the reader will find strongly contrasting perspectives on analyzing syntax, as they surface, for example, in the chapters by Boeckx on the one hand and O'Grady on the other, or of accounting for typological data, as they can be found, for example, in the chapters by Baker and Van Valin on the one hand and that of Haspelmath on the other.

In accordance with the general theme of this volume, authors tend to emphasize what is common to human languages across genetic and geographical boundaries and how the commonalities are best to be accounted for in linguistic analysis. The editors considered it important, however, to also draw the reader's attention to areas where languages differ, and they decided to devote one chapter to linguistic relativity and the effects it might have on purportedly non-linguistic cognition (Pederson, this volume).

In concluding, we wish to make it clear that a student looking for guidance on how to analyze a given language may be disappointed when consulting this book since its concern is not primarily with offering means of analysis but rather with a survey of "models" that may be of help in finding or developing the right framework for analyzing a language or set of linguistic data.

1.1 Goals

The main goal of this volume thus is to provide the student of language with alternatives that have been proposed in contemporary linguistics for analyzing and understanding the structure of human languages. To this end, the authors were confronted with the following questions that were meant to provide guidelines in the preparation of chapters:

(a) How can the main goals of your model be summarized?
(b) What are the central questions that linguistic science should pursue in the study of language?
(c) What kinds of categories are distinguished?
(d) What is the relation between lexicon, morphology, syntax, semantics, pragmatics, and phonology?
(e) How is the interaction between cognition and grammar defined?
(f) What counts as evidence in your model?
(g) How does your model account for typological diversity and universal features of human languages?
(h) How is the distinction synchrony vs. diachrony dealt with?
(i) Does your model take sociolinguistic phenomena into account?
(j) How does your model relate to studies of acquisition and to learning theory?
(k) How does your model generally relate to variation?
(l) How does your model deal with usage data?
(m) What kind of explanations does your model offer?
(n) How does your model relate to alternative models?

For good reasons, the authors of this volume highlight the potential that their work offers to the student of language or linguistics, and are therefore less concerned with areas where their work offers less satisfactory or no solutions. Accordingly, the way and the extent to which the questions are addressed in the following chapters differ greatly from one chapter to another. To be sure, not all of the questions are relevant to what a particular framework of linguistic analysis is about; hence, such questions are ignored by the authors concerned. There are also authors who relate their framework explicitly to the catalogue of questions, and simply admit that their work has scope only over a limited set of linguistic phenomena, or that it does not provide meaningful answers to specific questions. For example, Hudson (this volume) admits that his theory of Word Grammar has research gaps in areas such as phonology, language change, metaphor, and typology; or Van Valin (this volume) observes that there is no theory of phonology related to Role and Reference Grammar, and that work on morphology is in its initial stages; and Baker (this volume) notes that extending the kind of questions that he is concerned with in Formal Generative Typology to the domains of phonology and the lexicon would simply be outside his expertise.

Some of the questions received relatively little attention. This applies in particular to the question of what counts as evidence in a given model. In approaches relying largely or entirely on quantitative data, though, such as corpus linguistics or probabilistic linguistics (Biber, Bod, this volume), there is a clear answer to this question;

thus, Biber says: "Considered within the larger context of quantitative social science research, the major strengths of the corpus-based approach are its high reliability and external validity".

The main objective of this handbook is to have current influential approaches to linguistic analysis represented and to provide the reader with a convenient means of comparing and evaluating the various approaches. To this end, the editors aimed at reserving one chapter for each of the approaches. In a few cases, however, it turned out desirable to have more than one chapter devoted to one and the same approach in order to take account of contrasting orientations characterizing the relevant approach. Accordingly, Optimality Theory is represented with chapters on phonology (Gouskova), on the one hand, and on grammatical categories (de Swart and Zwarts), on the other, and the Chomskyan tradition of linguistics is represented not only with a general chapter on language-internal analysis (Boeckx) but also with chapters highlighting its potential of dealing with typological diversity (Baker) and of analyzing the cartography of syntactic structures (Cinque and Rizzi).

1.2 Approach, framework, model, program, or theory?

How to refer to one's work: Does it qualify as an "approach", a "framework", a "model", a "program", a "theory", or something else? The decisions made differ greatly from one author to another, depending on the goals underlying their work. Functional Discourse Grammar, Lexical-Functional Grammar, and others are theories (Asudeh and Toivonen, Hengeveld and Mackenzie, this volume), Minimalism is a program (Boeckx), natural semantic metalanguage is an approach (Goddard). But perhaps, more importantly, one and the same author may refer to his or her work as a "model" in some contexts, as an "approach" in other contexts, or as a "framework" in still other contexts. More generally, authors with a generativist orientation tend to phrase their work in terms of a theory, and, for equally good reasons, other linguists avoid this term; for quite a number of linguists with a functionalist orientation, there is some reluctance to recognize "theory" of any kind as being of use in doing linguistics.

The problem with the terminology is that there is not much agreement across the various schools on how to define these terms. Dryer (2006*a*: 28–9) says that "[t]he notion of theory widely assumed in formal linguistics is essentially equivalent to that of a metalanguage for describing languages. Providing an analysis of a particular set of data within a formal theory involves providing a description of that data within the metalanguage that constitutes that theory". But Haspelmath (this volume) uses a similar definition for "framework", characterized by him as a sophisticated and complex metalanguage for linguistic description that is intended to work for any language.

In addition to the diversity just outlined there is also some range of diversity in what should be the main goals of linguistic analysis. To give just a few examples, the main goal of Functional Discourse Grammar is to give an account of morphosyntactically and phonologically codified phenomena in languages (Hengeveld and Mackenzie, this volume), and Langacker (this volume) states that for Cognitive Grammar the goal is to describe the structure of particular languages and develop a general framework allowing the optimal description of any language. For others again, the declared goal is to explain the structure of language (e.g., Hudson, this volume) or to understand the cognitive organization of language (Bybee and Beckner, this volume).

1.3 Orientations

There is no shortage of classifications in the relevant literature proposing groupings and cleavages among the various approaches to linguistic analysis; the reader is referred to relevant works such as Newmeyer (1998), Darnell et al. (1999), Butler (2008), etc. for information. More generally, linguistic approaches tend to be divided into generativist (or formalist) and functionalist ones, and, when we submitted a proposal for the present book to the publisher, an anonymous reviewer suggested that "[s]omething could be said about the mutual antipathy that seems to exist between the two camps. Some formalists are dismissive of the functionalist approach, and some functionalists have a variety of negative feelings about formalism including in some cases an almost pathological fear of even drawing a tree diagram". An obvious way of structuring this volume might therefore have been to present the chapters in accordance with such a divide.

Our reasons for not adopting such a procedure are the following. First, we believe that the overall goals of linguists are essentially the same, namely understanding what languages are about, and how they can best be described and explained. Second, our main concern is not with differences in linguistic methodology but rather with finding answers to the questions listed above. And third, it would seem that this "divide" is gradually losing much of the significance it once had.

There is neither agreement on how the two main kinds of approaches should be referred to nor on what their distinguishing properties are. For reasons given in Newmeyer (1998: 7–11) we will refer to what are frequently called formal or formalist approaches as *generativist* ones and retain the widely accepted term *functionalist* for the second kind of approaches or orientation,[1] even if we do not know if the term "generativist" would really be accepted as a synonym for "formal", for example, by those doing work in representational frameworks outside of the Chomskyan Principles and Parameters or Minimalist traditions. The term "functionalist" is seemingly less controversial for the second kind of approaches, but the problems with this term are of a different nature: What is commonly subsumed under this label includes such a wide range of directions and schools that some feel tempted to subsume anything that excludes a generativist perspective under this label (see also below).

This divide between two basic orientations is associated with a range of contrasting perspectives; it surfaces in a number of antagonisms that have been pointed out in the relevant literature. One distinguishing feature is that for many generativists—but by no means for all—a central task for linguists is to characterize the formal relationships among grammatical elements largely independent of some characterization of the semantic and pragmatic properties of those elements. On the functionalist view, by contrast, language is explained with reference to the functions it is argued to serve. For most linguists with a functionalist orientation, language is foremost an instrument for communication, and the following claim made by Simon Dik (1986: 21) more than two decades ago is still endorsed by many students of language: "The primary aim of natural languages is the establishment of inter-human communication; other aims are either secondary or derived." This view contrasts with that dominant among linguists with a generativist orientation; for Chomsky (1980: 239), human language "is a system for free expression of thought, essentially independent of stimulus control, need-satisfaction or instrumental purpose".

In defense of the former view one might argue, as has in fact been done (e.g., Nuyts 1993), that many forms of presumed non-communicative behavior, such as self-talk, involve the same kind of mechanisms as linguistic communication and can be accounted for with reference to the latter. But in much the same way one can also argue that language is a tool of thought and that any linguistic communication presupposes thought, or cognition, and, hence, that communication is derivative of the latter.

Another area where contrasting opinions can be found is the following. On the one hand, there are approaches relying on the notion of Universal Grammar (UG) and an autonomous system of generative rules (e.g., Chomsky 1995); on the other hand, there are approaches that do without any form of universal grammar or the

[1] Note that names used to refer to approaches do not necessarily reflect the orientation concerned; for example, Lexical-Functional Grammar (see Asudeh and Toivonen, this volume) is not commonly considered to be a functionalist theory.

assumption that there is something like grammar and rules, arguing, as is done, for example, in the emergentist approach of O'Grady (this volume), that language evolves as a product of efficient processing. This raises the question of what the ontological status of a "rule" is or should be, whether rules are really required in linguistic analysis or, whether "rule-like behavior" may be no more than, for example, a side effect of maximizing probability, as is argued for by students of probabilistic linguistics (see Bod, this volume).

Another distinction concerns the question of whether linguistic analysis (and linguistic explanation) should focus on language knowledge or on language use. The central concern of generative grammar is with what constitutes knowledge of language, how this knowledge is acquired, and how it is put to use (Chomsky 1986*a*: 3). But there is also the contrasting view according to which the main concern of the linguist should be with understanding the structure of languages; as Bybee and Beckner (this volume) argue, "the units and structure of language emerge out of specific communicative events".

Finally, there is also the question relating to what should be the most central domain of linguistic analysis. In an attempt to summarize the main contrasting positions on this issue, Butler (2003*b*: 27) concludes that "a functional approach to language would place semantics/pragmatics at the very heart of the model, thus differing radically from formal approaches, which consider syntax as central". This is echoed, for example, in the conception of language in relevance theory (Yus, this volume), or of Systemic Functional Grammar, where language is viewed as meaning potential where all strata of the linguistic system contribute to the making of meaning (Caffarel, this volume). As the following chapters suggest, however, the answer to this question is not always all that unambiguous (see also below).

This is but a small catalog of distinguishing properties that have been mentioned. There are many other contrasting positions in addition; suffice it to mention that linguists working in the tradition of Michael Halliday emphasize that "[t]he image of language as rule is manifested in formal linguistics; the image of language as resource is manifested in functional linguistics" (Caffarel, this volume), and one might also mention that there is a remarkable pragmatic difference in the role played by the central exponents of the two "camps", characterized by Newmeyer thus:

> For better or worse [...], Chomsky is looked upon as the pied piper by the majority of generative linguists. No functionalist has managed to play the pipes nearly as enticingly to the graduate students of Hamlin. (Newmeyer 1998: 13)

That there is a fundamental divide in current linguistics on what language is, or is about, is undeniable, but this divide is far from watertight; rather, it is leaky and—as far as recent developments in general linguistics suggest—leakiness is increasing. First, neither students of approaches such as Relational Grammar, Lexical Functional Grammar, Head-Driven Phrase Structure Grammar, etc. nor many other students working on formal syntax would necessarily look upon Chomsky as the pied piper. Second, each of the two "camps" is associated with a wide range of different approaches and, as will become obvious in the following chapters, there are considerable areas of overlap between the two. As Van Valin (2000: 335–6) maintains, many of the ideas and methodologies subsumed under the heading "functional linguistics" are more distant from each other than they are from many formalist ideas. If one were to use his classification into purely formalist, structural-functionalist, and purely functionalist as a basis then Role and Reference Grammar is located in the intermediate category of generativist-functionalist approaches (Van Valin 1993*a*: 2). And according to Hengeveld and Mackenzie (this volume), Functional Discourse Grammar is located halfway between radical formal and radical functionalist approaches.

That degree of formalization is no significant distinguishing feature between generativist and functionalist approaches is also suggested by recent developments in what is commonly referred to as construction grammar: Whereas some directions within this general research paradigm, such as probabilistic linguistics (Bod, this volume) and Embodied Construction Grammar (Feldman et al., this volume), are highly formalist, others, such as radical construction grammar (Croft 2001) or cognitive grammar (Langacker, this volume) are distinctly less so.

Second, there are also commonalities among differing approaches. Many of the approaches, if not all, are concerned—in some way or other—with searching for the most appropriate, economic, or most elegant way of analyzing language structure. For example, when Langacker (this volume) concludes that Cognitive Grammar "shares with generative grammar the goal of explicitly describing language structure" then this also applies to many other approaches across the "divide". And all approaches have some typological basis, that is, they rest on generalizations about languages across genetic phyla and continents, even if the role played by typology varies considerably from one approach to another (see Baker, this volume, for discussion; see also below). And finally, in spite of all the specialized terminologies that characterize individual approaches, there is a common core of technical vocabulary figuring in many different theoretical frameworks. Terms such as sentence, verb, noun, determiner, agreement, passive, tense, aspect, negation, complement, voice, subordination, relative clause, etc. belong to the technical vocabulary of most approaches.

While linguists across different theoretical orientations in fact share a large range of technical vocabulary, one has to be aware, however, that there are also many contrasting definitions and uses of one and the same term, reflecting alternative theoretical orientations. When students of the natural semantic metalanguage approach discuss issues of "universal grammar" and "language universals" (Goddard, this volume) then the theoretical assumptions underlying this usage are fairly different from those that students working in the Chomskyan tradition make (see, for example, Baker, this volume). And much the same applies to a number of other terms; what is defined in Lexical Functional Grammar theory as a "functional constraint" (cf. Asudeh and Toivonen, this volume) has little in common with the use of the same term in many functionalist approaches. Conversely, there are also quite a number of cases where one and the same general linguistic phenomenon is referred to by different terms in the various approaches—what is called a "head" in many schools of linguistics corresponds to the "regent" in dependency theory, and the notion "subcategorization" corresponds in a number of ways to what in other traditions would be referred to as "valency" (Ágel and Fischer, this volume). Such differences are far from arbitrary; rather, they are indicative of the diversity of theoretical concepts that are the subject matter of the following chapters.

1.4 Locating linguistic analysis

There is consensus among most authors of this volume that linguistics is an autonomous discipline which requires its own theoretical foundation, methodological apparatus, and discipline-specific set of analytical techniques. Nevertheless, there are also authors arguing that linguistics is related to some other discipline in a principled way. Among the disciplines that are held to be particularly closely related to linguistics, especially in some more recent works, biology occupies a prominent position. Boeckx (this volume), for example, argues that "the generative enterprise is firmly grounded in biology" and, from a different perspective, Givón (this volume) emphasizes the analogies that exist between linguistics and biology; language diachrony, he argues, recapitulates many general features of biological evolution. Both abide by four principles of developmental control, namely graduality of change, adaptive motivation, terminal addition (of new structures to older ones), and local causation (with global consequences). Note also that Feldman et al. (this volume) claim that the correct linguistic analysis ultimately depends on evidence from biology, psychology, and other disciplines.

Up until the early 20th century, if not later, a common practice in linguistics was to use Latin grammar as a model for describing other languages, including non-European languages, and one of the major achievements of structuralism was that it freed linguistics of this straitjacket, making it possible to analyze each language in its own right. Now, it is argued by some modern linguists that after the 1950s the Latinist model was replaced in some schools of linguistics by an English model, in that the kinds of categorization used to describe grammatical structures to be found in the languages across the world were biased in favor of the categories of English; cf. Chomsky's (1981: 6) assertion that "[a] great deal can be learned about UG from the study of a single language". This is an issue that also surfaces in some of the chapters of this volume; Van Valin illustrates the problem with the following example:

> [...] theories starting from English and other familiar Indo-European languages often take the notion of subject for granted, whereas for one that starts from syntactically ergative and Philippine languages, this is not the case, and the notion of subject as a theoretical construct is called seriously into question. (Van Valin, this volume)

But then the question arises of what should be one's template or templates in deciding on how language structures should be analyzed. A somehow extreme perspective, one that is biased neither in favor of any theoretical presuppositions nor of some specific language or group of languages, is suggested by Haspelmath:

> The idea that a single uniform framework could be designed that naturally accommodates all languages is totally utopian at the moment. So instead of fitting a language into the Procrustean bed of an existing framework, we should liberate ourselves from the frameworks and describe languages in their own terms. (Haspelmath, this volume)

1.5 Analogies used for understanding linguistic phenomena

Metaphors and other analogical figures provide convenient means for demonstrating, illustrating, or understanding salient features of one's own road to linguistic analysis as against alternative roads. A paradigm example is provided by Newmeyer (1998: 161–2), who offers a couple of relevant analogies to describe the status of internal explanations for autonomous syntax. One relates to bodily organs, such as the liver, the other concerns the game of chess. Like the principles of generative syntax, he observes, those of chess form an autonomous system: Through a mechanical application of these principles, every "grammatical" game of chess can be generated. But he also observes that the autonomy of this game does not exclude the possibility that aspects of the system were motivated functionally. One kind of functional motivation can be seen in the aims of the original developers and the influence that players may have exerted on the rules of the game; another one concerns the players who, subject to the rules of the game, have free choice of which pieces to choose and where to move them. Nevertheless, such factors are irrelevant to the autonomy of chess, and he concludes:

> By the same reasoning, the autonomy of syntax is not challenged by the fact that external factors may have affected the grammar of some language or by the fact that a speaker of a language can choose what to say at a particular time. The only issue, as far as the autonomy of syntax is concerned, is whether one's syntactic competence incorporates such external

motivating factors. As we have seen, it does not do so. In short, the autonomy of syntax maintains that as a synchronic system, grammatical principles have an internal algebra. This fact, however, does not exclude the possibility that pressure from outside the system might lead to a changed internal algebra. (Newmeyer 1998: 161)

One may add that this does not conclude the list of analogical features shared by the two kinds of phenomena compared; for example, like chess, language is a social institution, created by humans for humans. Newmeyer's primary concern is with the system and the "internal algebra" of the principles underlying the system, including the competence of the persons concerned. But there are a number of alternative perspectives that one may adopt in analyzing such institutions, and each of these perspectives is associated with a different set of questions. Two possible alternatives are hinted at by Newmeyer. One of them would invite questions such as the following: Who designed the institution, and why was it designed in the first place? How, or to what extent, does the present design of the institution reflect the motivations of those who designed it? The other perspective would concern questions such as the following: What do people do with the institution? What are the aims and purposes for using it? And under what circumstances do they use it or not use it?

Such questions suggest that there are at least two contrasting ways of analyzing such institutions: One may either highlight their internal structure, the principles on which they are based, and the knowledge that people have about these institutions, or one may focus on those who developed and/or use these institutions. The latter perspective is found especially but not only among those following a functionalist orientation. The analogies favored by such scholars are of a different nature: Rather than games, body parts, or products of "architecture" (cf. Jackendoff, Culicover, this volume), they use analogies highlighting the role of the language user or processor, who may be likened to an architect or builder, as reflected in Hagège's (1993) metaphor of the "language builder", or a craftsman, as in O'Grady's emergentist framework (2005; this volume), where the native speaker is portrayed as a "language carpenter" designing sentences by combining lexical items.

Perhaps the most common metaphorical vehicle drawn on in linguistics is that of a biological phenomenon, namely that of trees. Both in diachronic and synchronic linguistics, the tree has provided a convenient template for describing and understanding taxonomic relationships. Throughout the history of linguistics, tree diagrams have been recruited to represent patterns of genetic relationship among languages, syntactic structures, and other phenomena. One issue that has found some attention in more recent discussions, reflected in the present volume, is whether or not tree branchings should necessarily be binary. But there are also authors doing without tree models; Hudson (this volume), for example, prefers to represent syntactic sentence structure as a network rather than in terms of any kind of a tree structure.

1.6 Domains of language structure

Roughly speaking, it would be possible to classify approaches on the basis of which domain or domains of language structure they are most centrally concerned with. But one question here is which domains are to be distinguished in the first place. The ones most commonly appearing in the following chapters are phonology, morphology, syntax, semantics, the lexicon, and pragmatics, but not all of them are recognized by all scholars as being significant domains of grammar.

One of the domains that has attracted the interest of linguists in the course of the last fifty years, perhaps more than others, is syntax. But should syntax be given a privileged status in analyzing language structure, as it is, for example, in the Minimalist Program and other approaches framed in the Chomskyan tradition, or should it be seen as functioning "in the grammar not as the fundamental generative mechanism, but rather as an intermediate stage in the mapping between meaning and sound" (Jackendoff, this volume), or as being derivative of cognition, as is argued for example in some models of cognitive grammar (see, for example, Langacker, this volume), or as being a product of discourse pragmatic forces, as is suggested by some functionalist linguists (cf. Givón 1979)? And should all syntactically sensitive phenomena of language structure—e.g., constituency on the one hand and grammatical functions on the other—be treated in one and the same domain, as is done in some of the syntactic approaches discussed in this volume, or should there be, for example, two separate structures (c-structure vs. f-structure), as students of Lexical-Functional Grammar propose (Asudeh and Toivonen, this volume)?

A number of students of grammar do recognize syntax as a distinct domain but do not attribute any central role to it. Rather than a piece of syntactic machinery, Givón (this volume) sees a well-coded lexicon together with some rudimentary combinatorial rules, as can be observed in pre-grammatical pidgin and other forms of communication, as more essential for understanding and analyzing language structure. In yet other directions of linguistics, syntax is not treated as a distinct domain of grammar at all. In particular, some linguists with a functionalist orientation argue that syntactic and morphological phenomena form an inextricable unit, referred to as "morphosyntax".

But, more so than syntax, morphology has been the subject of contrasting perspectives. In some schools of linguistics no distinct component or level of morphology is distinguished. Having at times been dubbed "the Poland of linguistics", some linguists do not consider morphology to be a relevant subdiscipline at all, treating it rather as a component of some other domain, whether that be syntax, phonology, or the lexicon. There is the view, for example, that morphology is included in the same comprehensive representational system as syntax (e.g., Baker, this volume), even if this view is not shared by many others. Spencer and Zwicky (1988*a*: 1), in contrast, view morphology as being at the conceptual center of linguistics. That morphology is a distinct domain and subdiscipline of linguistics is also maintained in other approaches discussed in this volume, such as Word Grammar (Hudson, this volume), and for yet others, morphology "is the grammar of a natural language at the word level", as Booij (this volume; see also Booij 2007) puts it.

And there is also a wide range of different opinions on the place of the lexicon in grammar. Some scholars would consider the lexicon to be only of marginal concern for analyzing grammar, or treat the lexicon and grammar as mutually exclusive phenomena (cf. Croft 2007: 339). Others again attribute core syntactic properties to the lexicon. According to adherents of dependency grammar and valency theory, an essential part of grammar is located in the lexicon, "in the potential of lexemes for connexion, junction, transfer and valency" (Ágel and Fischer, this volume; see also Hudson, this volume), and Categorial Grammar is, as Morrill puts it, "highly lexicalist; in the ideal case, purely lexicalist." Passivization is widely held to be an operation to be located in syntax; but in Lexical-Functional Grammar (LFG) or Head-Driven Phrase Structure Grammar (HPSG) it is treated as a rule that converts active verbs into passive verbs in the lexicon, altering their argument structure (see Jackendoff, this volume, for discussion). That the lexicon is of central importance for understanding and analyzing grammar is also pointed out in some other chapters of this volume, such as those of Givón and O'Grady.

Semantics is recognized as a distinct domain in most approaches, and in some approaches it is viewed as being the domain that is most central to linguistic analysis (Goddard, Caffarel this volume), even if it is looked at from a number of contrasting perspectives. That the meaning of both lexical items and constructions cannot be understood satisfactorily without adopting a frame-based analysis is argued for by Fillmore and Baker (this volume). Another theme concerns the place of semantics vis-à-vis other domains of linguistic analysis. There are different views on where semantics ends and other domains such as pragmatics or cognition begin. Huang (this volume) draws attention to Grice (1989), who had emphasized "the conceptual relation between natural meaning in the external world and non-natural, linguistic meaning of utterances"; we will return to this issue below. In some approaches there is an assumption to the effect that semantics is primary while pragmatics is secondary, the latter being concerned largely with phenomena that cannot be fitted into a semantic analysis. Jackendoff (this volume), by contrast, concludes that "one cannot do the 'semantics' first and paste in 'pragmatics' afterward", and in Role and Reference Grammar, discourse pragmatics plays an important role in the linking between syntax and semantics (Van Valin, this volume).

The boundary between semantics and pragmatics is in fact an issue that comes up in a number of chapters. Langacker (this volume) observes that the standard doctrine assumes a definite boundary between semantics and pragmatics (or between linguistic and extra-linguistic meaning) while he maintains that there is no specific line of demarcation between the two. A similar conclusion also surfaces in some lines of research in neo-Gricean pragmatics. Levinson (2000) argues that, contrary to Grice (1989), conversational implicatures can intrude upon truth-conditional content, and that one should reject the "received" view of the pragmatics–semantics interface, according to which the output of semantics provides input to pragmatics, which then maps literal meaning to speaker-meaning (Huang, this volume).

In a similar fashion, there is the question of what the place of pragmatics should be vis-à-vis syntax. Baker aptly portrays two main contrasting stances on this issue thus:

> On one view, pragmatics is the more basic study, and syntax is the crystallization (grammaticization) of pragmatic functions into more or less iconic grammatical forms. On the other view, syntactic principles determine what sentences can be formed, and then pragmatics takes the range of syntactic structures that are possible and assigns to each of them some natural pragmatic use(s) that take advantage of the grammatical forms that are available. The first view is characteristic of functionalist approaches to linguistics; the second is the traditional Chomskyan position. (Baker, this volume)

For a syntactic approach to deal with information structure, see Cinque and Rizzi (this volume); the role of pragmatics in the tradition of Grice (1978) is most pronounced in the chapters on Default Semantics (Jaszczolt, this volume), relevance theory (Yus, this volume), and on neo-Gricean pragmatic theory (Huang, this volume), and these chapters also show that much headway has been made in the research of this domain. Using a Kantian apophthegm of the form "pragmatics without syntax is empty; syntax without pragmatics is blind", Huang argues that pragmatics plays a crucial role in explaining many of the phenomena that are thought to be at the very heart of syntax.

Finally, there is the domain of cognition, which appears to be rapidly gaining importance in linguistic theorizing, and this is also reflected in some of the discussions of this volume. Boeckx (this volume), interprets the Minimalist Program of Chomsky (1995) as an "attempt to situate linguistic theory in the broader cognitive sciences", opening up fresh perspectives for an overall theory of cognition, and for Feldman et al. (this volume), linguistic analysis is part of a Unified Cognitive Science; they argue that the nature of human language and thought is heavily influenced by the neural circuitry that implements it, and integration of linguistic research with knowledge on neural reality is an important goal of their framework.

1.7 Relations among domains

If there are contrasting positions on which grammatical domains should be distinguished in linguistic analysis then this applies even more to the question of how the relationship among these domains should be defined: Are they all independent of one another, and, if they are not, how are they interrelated? It is this question where a particularly wide range of different answers is volunteered by the various authors. Jackendoff (this volume) argues that "the internal structure of some components of language, as well as the relation of language to other faculties, is consonant with a parallel architecture for language as a whole". In Cognitive Grammar, linguistic units are limited to semantic, phonological, and symbolic structures that are either part of occurring expressions or arise from them through abstraction and categorization, but Langacker (this volume) adds that a "major source of conceptual unification is the characterization of lexicon, morphology, and syntax as a continuum consisting solely in assemblies of symbolic structures".

Most authors state explicitly which domains they distinguish in their approach and what kinds of connections they postulate among domains. One way of establishing such connections is via hierarchical organization; Functional Discourse Grammar, for example, assumes a top-down organization of grammar, where pragmatics governs semantics, pragmatics and semantics govern morphosyntax, and pragmatics, semantics, and morphosyntax govern phonology (Hengeveld and Mackenzie, this volume). A tenet of some schools of linguistics is in fact that there is one domain that has a privileged status vis-à-vis other domains. Such a status can be due to the magnitude of connections that that domain is held to share with other domains. But such a status can also be due to relative degrees of descriptive and/or explanatory power attributed to one specific domain. In other approaches, specific theoretical devices are proposed to connect different domains. For example, students of Lexical-Functional Grammar use Glue Semantics as a theory to take care of the interface between syntax and semantics, and an "m-structure" is proposed to deal with the interface between syntax and morphology (Asudeh and Toivonen, this volume).

In yet other approaches there is some specific domain that relates different domains to one another. Thus, in Systemic Functional Grammar, a tristratal linguistic system of semantics, lexicogrammar, and phonology is proposed, and semantics is the interface between grammar and context (Caffarel, this volume). In Role and Reference Grammar, discourse pragmatics plays an important role in the linking between syntax and semantics; Van Valin proposes a linking algorithm that directly connects the semantic with the syntactic representation, and there is a direct mapping between the two representations. For Spencer and Zwicky (1988*a*: 1), by contrast, morphology is at the conceptual center of linguistics since it is the study of word structure, and words are at the interface of phonology, syntax, and semantics (see above).

Finally, there is also the position represented in the volume according to which there is no need to distinguish domains in the first place. In the natural semantic metalanguage approach, for example, it is argued that meaning is the bridge between language and cognition, and between language and culture, and Goddard (this volume) concludes that compartmentalizing language (or linguistic analysis) into syntax, morphology, semantics, and pragmatics therefore makes little sense.

1.8 The nature of structures

One issue discussed in a number of the chapters to be presented concerns the question of whether to set up a distinction between deep, underlying, or underived structures, on the one hand, and surface or derived structures, on the other, as is done in particular in approaches designed in the Chomskyan tradition or in Optimality Theory (Gouskova, this volume), or else whether such a distinction can or should be dispensed with, as is argued for in other, mostly but not only functionalist approaches (Hudson, Culicover, this volume).

A related question is how to deal with "zero", or empty categories, or null elements in syntactic representations, for example, with null pronouns (*pro*, PRO), noun phrase traces, "null subjects" in infinitival complements, or constructional null instantiation (Fillmore and Baker, this volume)—elements that are posited on the basis of structural considerations but are not phonologically expressed. Such categories have an important status for scholars working in some schools of syntactic analysis but are not recognized by others; Van Valin (this volume) describes the latter position thus: "If there's nothing there, there's nothing there" (see also Culicover, this volume).

Language structure shows both "regular" and "irregular" features; as Michaelis (this volume) puts it, many, if not most, of the grammatical facts that people appear to know cannot be resolved into general principles but must instead be stipulated. Linguistic approaches tend to highlight the "regular" structures, proposing generalizations that have a high degree of applicability. But what to do with the other part of grammar that is elusive to the generalizations, such as prefabricated word combinations or idiomatic and ritualized structures? This is a question that is addressed in some way or other in a number of chapters, and it is one where students of construction grammar and probabilistic linguistics propose answers challenging earlier models of grammar (Bod, this volume; see also Jackendoff, this volume, and others).

Another issue relates to the nature of grammatical categories. Most of the authors rely on entities of linguistic categorization that are discrete/algebraic, based on necessary and sufficient conditions, widely known as "classical categories". Others again, mainly but not only authors with a functionalist orientation, believe in the non-discreteness of linguistic categories, drawing on models framed in terms of Roschian prototypes (e.g., Langacker, Goddard, this volume) or of continuum models (see Taylor 1989 for a discussion of the distinguishing properties of these types of categories). Bod (this volume) observes that "[t]here is a growing realization that linguistic phenomena at all levels of representation, from phonological and morphological alternations to syntactic well-formedness judgments, display properties of continua and show markedly gradient behavior", and that all evidence available points to a probabilistic language faculty.

All these positions are represented in the following chapters, but there are also authors who allow for both kinds of categories. For example, in the Conceptual Semantics of Jackendoff and Culicover (this volume), conditions other than necessary and sufficient ones are admitted, and this type of semantics is compatible with Wittgensteinian family resemblance categories ("cluster concepts"). Bybee and Beckner (this volume) argue that "the boundaries of many categories of grammar are difficult to distinguish, usually because change occurs over time in a gradual way, moving an element along a continuum from one category to another". Accordingly, students of grammaticalization have proposed category structures that take the form of clines or chains but are not necessarily restricted to non-discrete categories (Heine and Narrog, this volume; see also Hopper and Traugott 2003).

Some discussions in the volume also concern two contrasting principles of analyzing syntactic relations, commonly described in terms of the distinction constituency vs. dependency: Should one use a phrase structure model, as quite a number of authors do, or a dependency model, as others prefer, or should one use a combination of both? There is fairly wide agreement that phrasal categories in some form or other constitute an indispensable tool for analyzing grammatical structures. But are they really indispensable in linguistic analysis, or are there alternative kinds of categories in addition, perhaps categories that allow us to do away with phrasal structures? Some students of language would answer this question in the affirmative. In dependency grammar and valency theory (Ágel and Fischer, this volume), but also in the emergentist approach of O'Grady (this volume), for example, it is argument dependencies rather than phrasal categories that are central, and the latter do not have any independent status in computational analysis.

Rather than phrasal categories, corpus-driven research finds other kinds of grammatical units and relations to be central. As Biber (this volume) observes, the strictest form of corpus-driven analysis assumes only the existence of word forms. Lexical bundles are such units; they are defined as the multi-word sequences that recur most frequently and are distributed widely across different texts; in English conversation they include word sequences like *I don't know if* or *I just wanted to*. Note that—unlike formulaic expressions—most lexical bundles cut across phrasal or clausal boundaries, being "structurally incomplete"; they are not idiomatic in meaning, and their occurrence is much more frequent than that of formulaic expressions (cf. also Haspelmath, this volume, who rejects notions such as noun phrase (NP) and verb phrase (VP), positing language-specific categories instead).

1.9 Lexical vs. functional categories

Another issue concerns the lexicon–grammar interface. Most approaches of linguistic analysis assume that in addition to lexical categories there are second kinds of forms, referred to as functional categories (or grammatical categories in some frameworks), operators, etc., i.e., grammatical taxa serving the expression of functions such as case, number, tense, aspect, negation, etc. The following are a few distinguishing properties that are widely recognized: (a) Lexical categories are open-class items while functional ones are closed-class items (having a severely restricted number of members belonging to the same class), (b) the former have a rich (lexical) meaning while that of functional categories is schematic, (c) lexical categories are independent words or roots while functional categories tend to be dependent elements, typically—though not necessarily—described as clitics or affixes, and (d) functional categories tend to be shorter (frequently monosyllabic).

In many perspectives this distinction is a robust one, but the way it is treated differs from one approach to another. There is in particular the question of whether the boundary between the two kinds of categories is discrete, as is maintained in the majority of approaches presented in this book, or gradual, as argued for explicitly in some of the approaches, or else, whether there is no boundary in the first place. Langacker (this volume), for example, maintains that instead of being dichotomous, lexicon and grammar form a continuum of meaningful structures, the primary difference between "lexical" and "grammatical" units being that the latter are more schematic in their content, their main import residing in construal.[2] A related position is maintained in grammaticalization theory, where it is argued that—at least in a number of cases—it is not possible to trace a discrete boundary between the two (see Heine and Narrog, this volume).

1.10 Recurring topics

As we observed above, there is a range of structural concepts and technical terms that are shared by most linguists. But at a closer look it turns out that there are also some dramatic differences on whether or how these concepts and terms apply within a given approach. For example, a paradigm concept of linguistic analysis across linguistic schools can be seen in the notion "sentence" (or clause). But

[2] Thus, Langacker argues that morphologically contrasting items such as nouns and derivational items like nominalizing suffixes can in a given case be assigned to the same category: "A nominalizer (like the ending on *complainer* or *explosion*) is itself a schematic noun and derives a noun from the verb stem it combines with" (Langacker, this volume).

there are also alternative notions that are proposed in some of the approaches. In approaches based on dependency theory (e.g., Ágel and Fischer, Hudson, this volume), the word is the basis of grammatical analysis, while Functional Discourse Grammar takes the discourse act rather than the sentence as its basic unit of analysis (Hengeveld and Mackenzie, this volume), and for a number of linguists "construction" is taken to be a crucial component of language structure (e.g., Michaelis, this volume). Some authors suggest that, rather than being epi-phenomenal, constructions should be in the center of linguistic analysis. To this end, a model is proposed in usage-based theory where the grammar of a language is interpreted as a collection of constructions organized into networks by the same criteria that words are (Bybee and Beckner, this volume). So, what should be the basis of linguistic analysis—words, sentences, constructions, discourse acts, or any combination of these? As we mentioned above, grammatical classes and syntactic structures other than word forms have no a priori status in corpus-driven research (Biber, this volume).

Another example concerns the case functions subject and object. That "subject" and "object" are useful or even indispensable entities for describing relations of arguments within the clause is widely acknowledged in many frameworks of linguistic analysis, even if various refinements have been proposed, such as that between the subject of a transitive clause (A) and that of an intransitive clause (S). However, the crosslinguistic validity of these entities has not gone unchallenged; suffice it to mention Van Valin's position (this volume) according to which grammatical relations like subject and direct object are not universal and cannot be taken as the basis for adequate grammatical theories (see above).

One of the grammatical phenomena that is seen by some to be a testing ground for the viability of a given approach is passivization. Questions that are raised in this volume on the analysis of passives include the following: Does passivization involve some movement of arguments, as is argued for in mainstream generative grammar, or is there no need to assume that there is movement, as maintained in a number of approaches, including Cognitive Grammar, Simpler Syntax, Role and Reference Grammar, etc. (see, for example, Langacker, Culicover, Van Valin, this volume)? And, is passivization really a syntactic phenomenon, as proposed in many approaches, or should it be treated as belonging to a distinct domain of information structure? Other phenomena that are discussed controversially in the following chapters are not hard to come by; one may wish to mention, for example, anaphora, which is analyzed by many as a syntactic phenomenon (cf. Chomsky 1995) but as a pragmatic one by others (see Huang, this volume).

Finally, the arrangement of linear elements in the clause has received remarkable attention across many schools of linguistics ever since Greenberg (1963*b*) proposed his classification of basic word order types. And once again, we find contrasting positions on what the ontological status of word order should be in a theory of language. There are, on the one hand, those who maintain that linear order must be a primitive of the syntactic theory rather than a derived property (Culicover, this volume; Barss and Lasnik 1986; Jackendoff 1990*b*). On the other hand, there are also those for whom linear order is derivative of other syntactic phenomena (Larson 1988), and for a number of functionalists word order is not a primitive of any kind but rather an epi-phenomenal product of discourse-pragmatic manipulation.

1.11 Typology

As we observed above, all approaches discussed in this volume rest at least to some extent on generalizations about different languages. But the role played by typology varies considerably from one approach to another. While accounting for typological diversity is a central concern for many authors (see, for example, Baker, de Swart and Zwarts, Hengeveld and Mackenzie, Haspelmath, Van Valin, this volume), this goal is not given high priority in the work of some other authors. One issue that is treated differentially is how and where typological distinctions should be accounted for in a given approach. For a number of authors, mostly of a functionalist orientation, typology forms the basis for all generalizations on language structure; in the Principles and Parameters model of Chomsky, by contrast, Universal Grammar (UG) is conceived as a set of principles regulating the shape of all languages, and these principles can be thought of as laws to which all languages must abide (see Boeckx, this volume); but in addition there is a set of parameters giving rise to the specific forms of individual languages, thereby accounting for typological diversity. Other, related, issues concern the level of abstraction that one should aim at in typological analysis and comparison, or the question of how many languages should be included in one's sample in order to come up with meaningful generalizations about the world's languages at large (see Baker, this volume, for an insightful discussion of this issue).

In a number of chapters, correlations are proposed between language typology and structural properties of the languages concerned. A case in point is provided by Huang (this volume), who observes that there is a correlation in his binding condition A between English-type, syntactic languages, where it is grammatically constructed, and Chinese-type, pragmatic languages, where it is pragmatically specified (for more observations of this kind, see de Swart and Zwarts, this volume).

The question of how grammatical categories relate to typological diversity and universal features of human languages is also an issue that surfaces in a number of the chapters. The answer given by most authors is that there should be a set of categories that is crosslinguistically the same, i.e., one that in some way or other reflects universal properties of human language or languages. For example, in the cartographic approach it is assumed that all languages share the same principles of phrase and clause composition and the same functional make-up of the clause and its phrases (Cinque and Rizzi, this volume). But there are also those authors who argue against universally uniform categories.

The latter position is associated, on the one hand, with a research field that has a long tradition in linguistics, namely the study of linguistic relativity. As the discussion by Pederson (this volume) suggests, some progress has been made more recently in our understanding of the interactions between cognitively universal and linguistically specific phenomena. On the other hand, it is also associated with another tradition of descriptive linguistics of the mid-20th century, namely American structural linguistics. Observing that the vast majority of the world's languages have not, or not sufficiently, been described and that the typologists' comparative concepts do not necessarily match the descriptive categories of individual languages, Haspelmath (this volume) suggests that language documentation is one of the primary tasks of the linguist and that one should describe each language and its grammatical categories in its own terms. With this view he takes issue not only with generativist approaches but also with those falling under the rubric of Basic Linguistic Theory, in particular with those of Dixon (1997) and Dryer (2006*b*), who use the same concepts for both description and comparison.

There is a stance that is somehow intermediate between these two extreme positions—one that is well represented in this volume—which postulates a crosslinguistically stable set of categories, but neither are all these categories represented in a given language nor are they represented the same way across languages.

A question central to all typological work is what typology can tell us about language universals. Here again there are contrasting positions correlating with the generativist/functionalist divide. For example, reviewing the monumental *World Atlas of Language Structures* (Haspelmath et al. 2005), Baker (this volume) concludes that "standard typologists have looked hardest for universals in exactly those domains where generativists least expect to find them, and have hardly looked at all in those domains where generativists predict that they exist".

1.12 Synchrony vs. diachrony

An issue that goes perhaps somewhat unnoticed in works on linguistic analysis concerns how a given account of language knowledge or language use relates to time: Should a framework used to explain the nature of language structure be restricted to synchronic observations or should it also account for diachrony?

That our understanding of linguistic structure may benefit from adding a diachronic perspective is argued for in a number of chapters, most notably in those of Givón and Heine and Narrog, and in some of the approaches a separate component is proposed to deal with the synchrony–diachrony interface; cf. the principle of viability in valency theory (Ágel and Fischer, this volume).

A distinction commonly made across the different schools of linguistics is one between two contrasting subfields, conveniently referred to, respectively, as *synchronic* and *diachronic* linguistics. The former deals with linguistic phenomena as they are observed at some specific point in time, which typically includes the present, while the latter is concerned with how linguistic phenomena behave across time, that is, essentially with how languages change. But, at a closer look, this does not really seem to be an exhaustive classification. That the borderline between the two is somewhat problematic has been pointed out independently by many linguists. Langacker (this volume), for example, notes that, since entrenchment and conventionality are matters of degree, there is never a sharp distinction between synchrony and diachrony.

Languages constantly change; the English of today is no longer exactly what it used to be a few decades ago: There are now new lexical items and use patterns that were uncommon or non-existent twenty years ago. Strictly speaking therefore, a synchronic analysis should relate to one specific point in time that needs to be defined in the analysis concerned. As a matter of fact, however, this hardly ever happens; grammatical descriptions are as a rule silent on this issue, and they may therefore more appropriately be dubbed achronic rather than synchronic. The fact that languages constantly change has induced some students of language to argue that a rigidly synchronic analysis is not possible or feasible[3] and that linguistic analysis should be panchronic in orientation (Heine et al. 1991; Hagège 1993), that is, include the factor of time as part of the analytic framework. To our knowledge, however, there is so far no general theory of language that appropriately accounts for panchrony.

1.13 Sociolinguistic phenomena

Linguistic analysis in general is based predominantly on language-internal phenomena, and this is reflected in the present volume. Accordingly, findings on the interface between linguistic and extra-linguistic phenomena, such as social or

[3] For example, Hagège (1993: 232) maintains: "No strictly synchronic study of a human language is conceivable, given the constant reshaping work done by LBs [language builders] even when a language seems to have reached a state of equilibrium."

cultural ones, play a relatively minor role in most of the approaches discussed in this volume; Hudson (this volume) therefore aptly concludes that "sociolinguistics has otherwise had virtually no impact on theories of language structure". But there are some noteworthy exceptions. As the chapters by Biber, Caffarel, Hudson, and Pederson in particular show, general linguistics can benefit greatly from incorporating a sociolinguistic or a socio-cultural dimension, and, in the natural semantic metalanguage approach, the notion cultural script is proposed as an important analytic tool to account for culture-dependent crosslinguistic distinctions, in particular on how ethnopragmatically defined categories can exert an influence on language structure, for example, in the form of constructions that—to use the wording of Goddard (this volume)—"are tailor-made to meet the communicative priorities of the culture".

1.14 On explanation

Differences surfacing in the volume also relate to how the structure of language or languages should be explained: Are explanations of the deductive-nomological type possible, meaningful, or both in linguistics, should explanations be context-dependent or context-free, should they be based on deduction, induction, abduction, or any combination of these, is it possible to draw a boundary between theory-internal and theory-external explanations, and can they be based on probabilistic generalizations, or is it only exceptionless, law-like generalizations that should be the concern of the linguist? And finally, should explanations be mono-causal or multi-causal, and should they be internal or external?

It is the last question that has attracted the attention of linguists perhaps more than others. In internal explanations, a set of facts falls out as a consequence of the deductive structure of a particular theory of grammar, or else a given phenomenon is explained with reference to other phenomena belonging to the same general domain. Thus, in approaches framed in the Chomskyan tradition, "one feature of a language is explained in terms of its similarity to another, at first different-seeming feature of that language and another language, by saying that both are consequences of the same general principle" (Baker, this volume); for example, Bresnan and Mchombo (1995) argue that Lexical Integrity provides a principled explanation of the complex syntactic, morphological, and prosodic properties of Bantu noun class markers (see Asudeh and Toivonen, this volume).

In external explanations, by contrast, a set of facts is derived as a consequence of facts or principles outside the domain of grammar. There is a partial correlation between these two types of explanations and the two main orientations of contemporary linguistics: Internal explanations are most likely to be found in works of generativist linguists, while functionalist approaches are invariably associated with external explanations. As will become apparent in the following chapters, however, the situation is much more complex. Another correlation concerns a distinction that we mentioned above, namely that between synchrony and diachrony: Internal explanations are overwhelmingly, though not necessarily, synchronic in orientation, while external ones are likely to have a diachronic component, or to be generally diachronic in nature.[4]

There is no overall way of deciding on which of the two kinds of explanations is to be preferred. As we saw above in our example of the game of chess, the nature of an explanation depends crucially on the kinds of questions one wishes to answer. Thus, whereas question (a) below is most strongly associated with internal explanations and generative theories, (b) is the question that functionalists tend to be concerned with, drawing on external explanations:

(a) How can the knowledge that native speakers have about their language best be understood and described?
(b) Why are languages structured and used the way they are?

Independently of how one wishes to decide on which of these options is to be preferred, there remains the question of whether there are certain domains, or components, of language structure that are particularly rewarding in looking for explanations. Syntax is the domain that has attracted quite some attention but—as we saw above—there is also the opinion, favored in functionalist traditions, that it is semantics or the pragmatics of communication that provide the best access to linguistic explanation.

1.15 Language evolution

A few decades ago, topics relating to language origin and language evolution were essentially non-issues in mainstream linguistics. The present volume suggests that this situation may have changed: Language evolution is nowadays a hotly contested subject matter in some schools of linguistics, both in functionalist works (Givón, this volume) and generativist traditions of modern linguistics (Jackendoff 2002; Hauser, Chomsky, and Fitch 2002; Jackendoff and Pinker 2005*a*). Givón

[4] A number of functionalists would follow Givón (1979; 2005) in arguing that, like functional explanation in biology, functional explanation in linguistics is necessarily diachronic (Bybee 1988*c*; Keller 1994; Haspelmath 1999*a*; 2008; Bybee and Beckner, this volume); in the words of Dryer (2006*a*: 56), "a theory of why languages are the way they are is fundamentally a theory of language change".

in particular argues that relating language structure to other human and non-human communication systems and to language evolution is an essential task of understanding grammar (Givón 1979; 2005). A field of research that is held to be particularly rewarding in the reconstruction of language evolution is that of signed languages (see Wilcox and Wilcox, this volume). While there are contrasting opinions on how language evolution may have proceeded, there is agreement among these authors that it is essentially possible to reconstruct this evolution.

1.16 Conclusions

For obvious reasons, the 33 chapters of this handbook are restricted to a specific goal and, hence, to a specific range of subject matters that are immediately relevant to the analysis of a given language, or of language in general. Accordingly, many other issues that are also of interest to the student of language have to be ignored. Such issues concern, on the one hand, comparative linguistics; as desirable as it would have been to have a more extensive treatment of relationship among languages, this would have been beyond the scope of the present volume. Other issues that will not receive much attention either are language contact, that is, how speakers of different languages and dialects interact with and influence one another, language internal diversity, such as variation among dialects or socially defined linguistic varieties, and languages that are said to enjoy a special sociolinguistic or linguistic status, such as pidgins and creoles. And finally, we also ask for the reader's understanding that many other subjects, such as language acquisition, applied linguistics, writing systems, language planning, clinical and forensic linguistics, to name but a few, could not be covered in the present volume.

This page intentionally left blank

CHAPTER 2

THE ADAPTIVE APPROACH TO GRAMMAR

T. GIVÓN

2.1 General orientation

Language is, by all accounts, one of the defining characteristics of *homo sapiens*. It is deployed in a wide range of adaptive contexts: social interaction, cultural transmission, education, literature, theater, music, humor and play, love and war. Of this rich array of useful applications, one may single out two core adaptive functions that make all the rest possible: (i) the *mental representation*, and (ii) the *communication* of information.

Mental representation is the affair of a single mind striving to code, make sense of, and in a way actively construct "reality", be it external, mental, or social. Communication is an affair of two minds exchanging mentally represented information. Of these two core functions, mental representation is both developmentally older and logically prior. One can represent information in the mind/brain without necessarily intending to communicate it; but one cannot communicate information that is not first represented in the mind.

Mental representation is as old as biological organisms. At whatever lever of complexity, organisms depend for their survival on sorting out the myriad individual *tokens* of incoming experience into much fewer adaptively-significant *types* (categories), and then tailoring their reaction (behavior) to the adaptive value of those categories.

Communication is as old as social organisms and perhaps even older in one important domain—sexual reproduction. All sexual, and even more so social, species communicate, at the very least within a restricted range of adaptively-relevant domains: Foraging, territorial control, social rank, aggression, mating, and rearing of the young (Cheyney and Seyfarth 2007). Pre-human communication seldom exceeds this range.

Homo sapiens, in diverging from the primate line, has expanded immensely the extent to which its survival depends on communication, doing so along two parallel tracks: (i) functionally, extending greatly the range of adaptive domains (topics) of communication; and (ii) structurally, creating a vastly more elaborate instrument of communication—language.

Biology has been an unabashedly adaptive-functional discipline ever since Aristotle, who in his *De Partibus Animalium* observed that in biological design, as in the design of man-made instruments, the exhaustive description of structures makes no sense unless use referred to their function (purpose). Put another way, biological design is driven by some *teleology*. The theoretical account of this teleology, the mystery of how structures and functions got to be paired, had to wait over two millennia for Darwin. Still, Aristotle remains the true father of the adaptive approach to the study of living organisms.

Somewhat paradoxically, Aristotle is also the father of an influential non-adaptive approach to linguistics, *structuralism*. In the opening paragraph of his *De Interpretatione*, Aristotle asserts that external reality is faithfully (iconically) mirrored by the mind; but that the mapping from mind to language is arbitrary. This latter observation was based on the diversity of sound sequences that can code roughly the same concepts (God = Deus = Allah = Watan Tanka = Núumaroghomapʉgat, etc.). In another of his books, *The Categories*, however, Aristotle implies a non-arbitrary (iconic) mapping between propositional logic (meaning) and grammar (form).

Ever since Aristotle, linguists and philosophers have been seesawing between a functionalist and structuralist view of language. Both positions were argued by Plato, in his *Cratylus* dialog. Towering figures in the history of structuralism in linguistics are F. de Saussure, L. Bloomfield, and N. Chomsky. Towering figures in the history of functionalism in linguistics are W. von Humboldt, H. Paul, E. Sapir, and O. Jespersen.

2.2 Representation and communication

The two core adaptive functions of human language are the representation and communication of information (knowledge, experience). We may take it for granted, given the overwhelming evidence from animal communication, child language development, and neurology (Geary 2005; Cheyney and Seyfarth 2007; Givon 2002, chs. 4, 5), that cognitive representation preceded communication in evolution, is present in pre-human species, and is a developmental prerequisite to language. What human communication added to the pre-existing cognitive representation system are two specific communicative codes—*phonology* and *grammar*.

Cognitive psychologists have long recognized three major systems of mental representation in the human mind/brain (Atkinson and Shiffrin 1968). The linguistic equivalents of these systems are sufficiently transparent.

(1) Major cognitive representational system:

Cognitive label	Linguistic equivalent
• permanent semantic memory	the mental lexicon
• episodic memory	the current text
• working memory and attention	the current speech situation

Not only are these three types of mental representation recognized for their specific cognitive-behavioral properties but also for their specific brain locations. In the next sections I will discuss the three briefly.

(a) Semantic memory

Semantic memory is the *mental lexicon*, a long-term repository of relatively stable concepts of *types* of entities (nouns), states or qualities (adjectives), or events (verbs). It is thus the repository of the our culturally shared view of external, mental, and social world. The mental lexicon is most likely organized as a network of conceptual nodes and connections (Givón 2005, chs. 2, 3), within which semantically-related word-nodes automatically activate each other (spreading activation; Swinney 1979). In addition to the more abstract core of this semantic network in the left pre-frontal cortex (Posner and Pavese 1997; Abdulaev and Posner 1997), more concrete perceptual and affective brain loci are also automatically activated by words with concrete—visual, auditory, olfactory, savory, tactile, or affective—meanings (Caramazza and Mahon 2006). Semantic memory is *cross-modal* (linguistic, visual, auditory, etc.; see Humphrey and Riddoch eds. 1987), and it is likely that language was but the latest evolutionary addition to its pre-human, pre-linguistic precursor (Givón 2002, chs. 4, 5).

(b) Episodic memory

Episodic (declarative) memory is the long-term repository of *propositional information* about unique events, states, or specific individuals, all known to us through lifelong experience; or of their concatenations in longer chunks of coherent discourse (Kintsch and van Dijk 1978; Gernsbacher 1990; Ericsson and Kintsch 1997; Givón

1995, ch. 8). Information comes into episodic memory via either non-linguistic (perceptual) or linguistic channels, and is then kept first in a temporary malleable sub-cortical processor (*hippocampus* and *amygdala*; Squire 1987; Petri and Mishkin 1994; Ericsson and Kintsch 1997). Information that merits longer-term, more stable representation is transferred later to a frontal-cortical locus (Squire 1987).

(c) Working memory and attention

Working (short term) memory represents what is available for immediate attentional activation. It thus overlaps partially with the attentional system (Schneider and Chein 2003; Posner and Fan 2004). Working memory is a limited storage-and-processing buffer of small capacity and short duration, where material is kept temporarily pending further processing options. It has a cross-modal conscious component that interacts with the *executive attention* (Gathercole and Baddeley 1993; Schneider and Chein 2003; Posner and Fan 2004), as well as several modality-specific non-conscious components (visual, auditory, tactile, etc; Gathercole and Baddeley 1993). In language processing, working memory is an important buffer where short chunks of information are represented verbatim, pending further processing decisions.

2.3 Human language as a combinatorial system

The most well-entrenched idea about the function of grammar, long licensed by linguists and adopted uncritically by others, is that grammar is a set of rules that govern the combination of words and morphemes into propositions (clauses). This misperception about grammar's adaptive niche in human communication is only natural, given two ubiquitous habits of linguists: (i) a methodology that inspects clauses/propositions in isolation from natural communicative context, and is thus dependent on Chomsky's (1965) notion of *competence*; and (ii) a theoretical perspective that emphasizes *event frames* (argument structure) at the expense of multi-propositional coherence. The most cogent articulation of these habits may be seen in Chomsky (1965, ch. 2), where *deep structure* (event frames) receive a coherent functional characterization (logical-semantic structure).

Chomsky's deep structure turns out to be the most common, and semantically most transparent, type of syntactic structures in natural communication: the *main, declarative, affirmative, active* clause. This clause-type is rightly recognized as the foundation of our study of the combinatorial nature of propositions, couched in terms of *phrase structure rules*. On the other hand, the function of the much more numerous "variant" syntactic structures—*transformed surface structures*—was left unmotivated, or dismissed as "stylistic" in Chomsky's early work (1965, ch. 3). But it is in studying this much larger set of syntactic structures that one finds the clues to the adaptive (communicative) function of grammar.

Chomsky's distinction between simple (deep structure = unmarked) and complex (transformed = marked) clause remains fundamental to our understanding of syntax. One may thus line up syntactic clause-types as follows (Givón 1995, 2001*a*):

(2)

Simple (unmarked)	Complex (marked)	Typical examples
main	subordinate	REL-clause, V-comp, ADV-clause
declarative	non-declarative	imperative, interrogative
affirmative	negative	negative
active-transitive	de-transitive	passive, antipassive, inverse
default topic/focus	marked topic/focus	L-dislocation, cleft, Y-movement

In terms of usage frequency in face-to-face communication, the simple clause-type is statistically predominant, at the level of 90%–95%, underscoring its privileged informational status. Not surprisingly, it is also cognitively easier to process (Givón 1995, ch. 2).

The combinatorial relation between lexical semantics, propositional information, and multi-propositional coherence, and the privileged role of grammar in constructing multi-propositional discourse, may be illustrated with the following simple example. Consider first the set of lexical words in (3) below:

(3) Lexical concepts (words):

a. eventually	b. police	c. conclude
d. dancer	e. drive	f. insane
g. director	h. proposition	i. lewdly
j. shoot	k. gun	l. smuggle
m. hall	n. night	o. before

We understand the meaning of these words regardless of the propositions in which they may be embedded in actual communication, presumably through some prototypical network activation in semantic memory (Swinney 1979).

With the addition of appropriate grammatical morphology (boldfaced below), we can combine them into coherent simple propositions, as in:

(4) Propositions (clauses):

a. Eventually **the** police conclud**ed that** [. . .]
b. [...] dr**o**ve **the** dancer insane.
c. **The** director proposition**ed the** dancer lewdly.
d. **The** dancer sh**o**t **the** director **with a** gun.
e. **The** dancer smuggl**ed the** gun **into the** hall the night before.

We understand the meaning of these "atomic" propositions, albeit in a somewhat "generic" way, regardless of the communicative context in which they may be embedded—provided of course that we understand the meaning of the component lexical words and function of the grammatical morphemes.

With proper adjustment of the grammatical morphology and the application of other syntactic devices (rules), we can also combine the five simple propositions in (4) into a coherent multi-propositional connected discourse, as in:

(5) Multi-propositional discourse:

a. Eventually **the** police conclud**ed that**,
b. **having been** driv**en** insane
c. **by the** director**'s** lewd proposition**ing**,
d. **the** dancer sh**o**t **him with a** gun
e. **which she had** smuggl**ed into the** hall **the** night before.

Now, if we were to re-order the connected propositions in (5) without adjusting their grammatical structure, the resulting discourse, as in (6) below, would be incoherent:

(6) c. **By the** director's lewd proposition**ing**
b. **having been** driv**en** insane
d. **the** dancer sh**o**t **him with a** gun
a. eventually **the** police conclud**ed that**
e. **which she had** smuggl**ed into the** hall **the** night before.

Some of the incoherence of (6) as a connected discourse is of course due to the new order itself: Events have their own real-world coherence. (Normally one aims and presses the trigger before one shoots the gun, and the victim falls dead only subsequently). But if we re-adjust the grammatical form of the clauses in (6), their re-ordered sequence may now yield a coherent—if different—discourse, as in:

(7) c. **Because he** proposition**ed her so** lewdly
b. and **thus** drove **her** insane,
d. **the** dancer sh**o**t **the** director **with a** gun, **which**,
a. **as the** police eventually conclud**ed**,
e. **had been** smuggl**ed** into the hall the night before **by the** dancer **herself**.

Not quite as elegant perhaps, but still coherent. As this simple-minded example demonstrates, it is the communicative coherence requirements of **multi-propositional discourse**, rather than the combinatorial semantic demands of atomic propositions *per se*, that motivate our use of the grammatical packaging of the same "deep structures". Functionally-oriented grammarians have been fond of saying that grammar is therefore determined by the **discourse context**. As I will try to suggest further below, "discourse context" is but a methodological heuristic stand-in for something else—the speaker's mental representation of the presumed mental states of the interlocutor during the ongoing communication.

2.4 Grammar

2.4.1 Preliminaries

Grammar is probably the latest evolutionary addition to the mechanisms that drive human communication (Givón 1979, 2002, 2005; Lieberman 1984; Bickerton 1981, 1990; Li 2002; Cheney and Seyfarth 2007). While the evolutionary argument remains necessarily conjectural, it is supported by a coherent body of suggestive evidence.

In ontogeny, children acquire the lexicon first, using it in pre-grammatical (pidgin) communication before acquiring grammar (Bloom 1973; Bowerman 1973; Scollon 1976; Givón 1979, 1990*b*; Bickerton 1981, 1990). Likewise, natural second language acquisition by adults follows a similar course, but without formal instruction it most commonly stops short of grammaticalization, remaining at the pidgin stage (Bickerton 1981, 1990; Bickerton and Odo 1976; Selinker 1972; Schumann 1978; Givón 1979, 1990*b*).

A well-coded lexicon can be acquired by many non-human species (Premack 1971; Gardner and Gardner 1971; Terrace 1985; Savage-Rumbaugh et al. 1993; Savage-Rumbaugh and Lewin 1994; Pepperberg 1999; Tomasello and Call 1997; Cheney and Seyfarth 2007; *inter alia*). This supports the suggestion that the neuro-cognitive structures that underlie semantic memory are old pre-human, pre-linguistic structures (Givón 2002, chs. 4, 5; Cheney and Seyfarth 2007).

In contrast, the natural communicative use of grammar in non-human species has never been attested. Nor has much success been reported in teaching grammar to non-human species (Premack 1971; Terrace 1985; Pepperberg 1999; Tomasello and Call 1997; Givón and Sue Savage-Rumbaugh 2006). Grammar as we know it seems to be a unique human capacity.

2.4.2 Grammar as structure

As a symbolic code, grammar is much more complex and abstract than the sensory-motor (phonological) apparatus that codes the lexicon. At its most concrete, the primary grammatical signal combines four major coding devices:[1]

(8) Primary grammar-coding devices:

- Morphology
- Intonation:
  - clause-level melodic contours
  - word-level stress or tone
- Rhythmics:
  - pace or length
  - pauses
- Sequential order of words or morphemes

Some of the primary coding devices (morphology, intonation) are more concrete, relying on the same sensory-motor devices that code the lexicon. But these concrete devices are integrated into a complex system with the more abstract elements (rhythmics, sequential order) that are no doubt second- or third-order constructs. The most concrete element of the grammatical code, grammatical morphology, is a diachronic derivative of lexical words (Givón 1971, 1979; Traugott and Heine (eds.) 1991*a*; Heine et al. 1991; Hopper and Traugott 1993; Bybee et al. 1994; *inter alia*).

The primary grammar-coding devices in (8) are in turn used to signal yet more abstract levels of grammatical organization:

(9) More abstract levels of grammar:

- Hierarchic constituency
- Grammatical relations (subject, object)
- Syntactic categories (noun, verb, adjective; noun phrase, verb phrase)
- Scope and relevance relations (operator-operand, noun-modifier, subject-predicate)
- Government and control relations (agreement, co-reference, finiteness)

The structural elements in (8) and (9) combine to create the various grammatical *constructions*, or *clause-types* (see (2) above). And it is those constructions, with their attendant morphology, that most directly code the various communicative functions.

[1] The first-order formal properties cited here are relatively concrete and perceptually accessible. More abstract approaches to syntax may reject some of those, including the entire notion of syntactic construction (Chomsky 1992), and may count other abstract properties not mentioned here.

2.4.3 Grammar as adaptive function

The adaptive function of grammar comes into sharp relief when one notes that humans can, in some developmental, social, or neurological contexts, communicate without grammar. In such contexts, we use the well-coded lexicon together with some rudimentary combinatorial rules. That is, we use *pre-grammatical pidgin* communication (Bloom 1973; Bowerman 1973; Scollon 1976; Bickerton 1981, 1990; Bickerton and Odo 1976; Selinker 1972; Schumann 1976, 1978, 1985; Andersen 1979; Givón 1979, 1990*b*).

The structural and functional differences between pre-grammatical pidgin and grammatical communication may be summarized as follows (Givón 1979, 1989):

(10) Pre-grammatical vs. grammatical communication

Properties	Grammatical	Pre-grammatical
STRUCTURAL:		
a. morphology	abundant	absent
b. constructions	complex, embedded, hierarchic	simple, conjoined, non-hierarchic
c. word-order	grammatical (subj/obj)	pragmatic (topic/comment)
d. pauses	fewer, shorter	copious, longer
FUNCTIONAL:		
e. processing speed	fast	slow
f. mental effort	effortless	laborious
g. error rate	lower	higher
h. context dependence	lower	higher
i. processing mode	automated	attended
j. development	later	earlier
k. consciousness	sub-conscious	more conscious

The heavy dependency of pidgin communication on the lexicon tallies with the fact that lexicon is acquired before grammar in both first and second language acquisition, as well as with the fact that more abstract vocabulary is the diachronic precursor of grammatical morphology in grammaticalization. Pre-grammatical children, adult pidgin speakers, and agrammatic aphasics comprehend and produce coherent multi-propositional discourse, albeit at slower speeds and higher error rates than those characteristic of grammatical communication. The identification of grammar with a more automated, subconscious, speeded-up processing system has been suggested in Givón (1979, 1989), Blumstein and Milberg (1983), and Lieberman (1984). Phonology, the other human communicative code, is likewise highly automated and subconscious.

2.5 Grammar and other minds

A context is a psychological construct.
(Sperber and Wilson 1986: 15)

Mind reading pervades language.
(Cheney and Seyfarth 2007: 244)

We noted earlier above that the adaptive function of grammar was to code the communicative (discourse) context of the proposition/clause. But we also noted that this notion of context-as-text was only a methodological heuristic. To begin with, "context" in cognition, and even more so in communication, is not an objective entity but rather a mental construct. Further, what the use of grammar is sensitive to, what grammar is adapted to do, is to represent—systematically, in the mind of the speaker-hearer—the constantly shifting epistemic and deontic states that the interlocutor is presumed to hold during ongoing communication. In other words, grammar is an adapted code for the mental representation of *other minds*, or what is currently known in cognitive neuro-science as *theory of mind*.

Communicating without a theory of mind is either implausible or inordinately slow, cumbersome and error prone, a message implicit in Grice's (1968/1975) influential paper on the pragmatics of communication. As Cheney and Seyfarth put it more recently (2007), "Mind reading pervades language". An extensive treatment of this subject may be found in Givón (2005). For the purpose of this chapter, a few illustrative examples will have to suffice.[2]

2.5.1 Mental models of epistemic states

The first example is taken from the grammar of referential coherence (reference tracking), a sub-function of the grammar that involves a huge number of constructions and grammatical morphology (Givón 2005, ch. 5). Consider the mid-discourse narrative in (11) below:

(11) a. There was **a** man standing near the bar,
b. but we ignored **him** and went on across the room,
c. where **another man** was playing the pinball machine.

[2] The literature on "theory of mind" is mind boggling, multi-disciplinary, and exponentially proliferating, going back to Premack and Woodruff's (1978) seminal contribution. For some of the discussion, see Baron-Cohen (2000); Byrne and Whiten (eds.) (1988); Cheney and Seyfarth (2007, ch. 10); Decety an Sommerville (2003); Decety and Jackson (2006); Givón (2005); Gopnik and Wellman (1992); Heyes (1998); Leslie and Frith (1988); Malle et al. (eds.) (2000); Melzoff and Prinz (eds. 2002); Povinelli and Preuss (1995); Tomasello et al. (2007); Wellman (1990); Whiten (ed.) (1991).

d. **We** sat down and ordered a beer.
e. **The bartender** took his time,
f. I guess **he** was busy.
g. So **we** just sat there waiting,
h. when all of a sudden **the man standing next to the bar** got up and screamed.

In marking "man", introduced for the first time in (11a), with the indefinite "a", the speaker cued the hearer that he doesn't expect him/her to have an *episodic-memory* trace of the referent. In coding the same referent with the anaphoric pronoun "him" in (11b), the speaker assumes that the referent is not only accessible but is still *currently activated*; that is, it is still under *focal attention*.

Another referent is introduced for the first time in (11c), this time with the indefinite marker "another". In using of the first-person pronoun "we" in (11d), next, the speaker assumes that his/her own referential identity is accessible to the hearer from the immediate speech situation, available in *working memory*. "The bartender" is introduced for the first time in (11e)—but marked as *definite*. This is so because the prior discourse had activated "bar", which then remained activated by the persistence of the narrated situation. And "bartender" is an automatically-activated connected node of the lexical frame "bar", thus a consequence of the cultural specificity of *semantic memory*. In continuing with the anaphoric pronoun "he" in (11f), the speaker again assumes that the referent is both accessible and currently activated, i.e., still under *focal attention*. And in using the first-person pronoun "we" in (11g), the speaker assumes that his own identity is accessible to the hearer in the *current speech situation*, still held in *working memory*.

Finally, the man introduced earlier in (11a, b) and then absent for five intervening clauses, is re-introduced in (11h). The use of a definite article suggests that the speaker assumes that this referent is still accessible in the hearer's *episodic memory* but that the hearer's memory search is not going to be simple. Another man has been mentioned in the intervening (11c) as "playing the pinball machine". Both referents are assumed to still be accessible in *episodic memory*, and would thus compete for the simple definite description "the man". To differentiate between the two, a *restrictive relative clause* is used, matching "standing next to the bar" in (11h) with the proposition "a man standing near the bar" in (11a). In using this grammatical cue, the speaker reveals his/her assumption that the hearer still has an episodic trace of both the referent and the proposition in (11a).

2.5.2 Mental models of deontic states

Example (11) above reveals another important feature of our presumption of access to other minds: Our mental models of the mind of the interlocutor shift constantly, from one clause to the next, during ongoing communication. As speakers release more information, they constantly update what they assume that the hearer knows; that is, the hearer's constantly shifting *epistemic* (knowledge) states. In this section we will see that speakers also possess running mental models of the hearer's constantly shifting *deontic* (intentional) states.

The deontic (and epistemic) states we will consider here are coded by the cluster of grammatical sub-systems that mark propositional modalities (Givón 2005, ch. 6). The most conspicuous of these sub-systems, and the easiest to illustrate, is the grammar of *speech acts*.

The study of speech acts has traditionally centered on a set of *felicity conditions* (*use conventions*) associated with declarative, imperative, interrogative, and other speech acts. These conventions have had an illustrious history in post-Wittgensteinean philosophy and linguistics (Austin 1962; Searle 1969; Cole and Morgan (eds.) 1975; *inter alia*). They are also known as *conventional implicature* (Grice 1975; Levinson 2000).

As an illustration, consider the following, somewhat schematic but still plausible, dialogue between speakers A and B:

(12)
A-i: So she got up and left.
B-i: You didn't stop her?
A-ii: Would you?
B-ii: I don't know. Where was she sitting?
A-iii: Why?
B-iii: Never mind, just tell me.

In the first conversational turn (12A-i), speaker A executes a *declarative* speech act, which involves, roughly, the following presuppositions about hearer B's current mental states (in addition to the speaker's own mental states):

(13) a. Speaker's belief about hearer's epistemic state:
- Speaker believes hearer doesn't know proposition (12A-i).
- Speaker believes hearer believes that speaker speaks with authority about proposition (12A-i).

b. Speaker's belief about hearer's deontic state:
- Speaker believes hearer is well-disposed toward the speaker communicating to him/her proposition (12A-i).

c. Speaker's own epistemic state:
- Speaker believes he/she knows in proposition (12A-i).

d. Speaker's own deontic state:
- Speaker intends to inform hearer of proposition (12A-i).

In the next turn (12B-i), B, the speaker, now executes an *interrogative* speech act (yes/no question), which involves, roughly, the following presuppositions about hearer A's current mental states (as well as the speaker's own):

(14) a. Speaker's belief about hearer's epistemic state:
- Speaker believes hearer knows the declarative proposition underlying question (12B-i).
- Speaker believes hearer knows speaker does not know that proposition.

b. Speaker's belief about hearer's deontic state:
- Speaker believes hearer is willing to share their knowledge of that proposition.

c. Speaker's own epistemic state:
- Speaker is not certain of the epistemic status of the proposition underlying (12B-i).

d. Speaker's own deontic state:
- Speaker would like hearer to share their knowledge with him/her.

In turn (12Biii), lastly, speaker B executes a *manipulative* speech act, which involves, roughly, the following presuppositions about hearer A's current mental states (as well as the speaker's own):

(15) a. Speaker's belief about hearer's epistemic state:
- The hearer believes the hearer knows that the desired event ("You tell me") is yet unrealized.

b. Speaker's belief about hearer's deontic state:
- Speaker believes hearer is capable of acting so as to bring about the desired event.
- Speaker believes the hearer is well-disposed toward acting to bring about the desired event.

c. Speaker's own epistemic state:
- Speaker believes the desired event ("You tell me") is yet unrealized.

d. Speaker's own deontic state:
- Speaker would like the event ("You tell me") to come about.

At every new turn in conversation (12), not only do the speaker's own belief-and-intention states change, but also his/her mental representation of the hearer's belief-and-intention states. And one would assume that a similar fast-paced adjustment also occurs in the hearer's mental model of the speaker's belief-and-intention states.

2.6 Development, change, and evolution

2.6.1 Parallelisms

In biology one takes it for granted that the way extant organisms are today, and the current distribution of their staggering variety, is most cogently explained by the two *developmental processes* that brought the organism(s) up to this point—*ontogenesis* (embryology) and *phylogenesis* (evolution). This is how Darwin explained the cumulative descriptive and classificatory data gleaned during the preceding 2,000-odd years of systematic study of biology, from Aristotle to Linaeus.

In linguistics, in a rather striking analogy, three major developmental trends have jointly fashioned the way language—and languages—are now:

- *Evolution*: The descent of the language capacity of the human species.
- *Ontogenesis*: The emergence of language in children.
- *Diachrony*: The historical development of particular languages.

Of the three, diachrony is a uniquely human phenomenon, a cumulative historical accretion that is largely unattested elsewhere in biology. Diachrony has the most direct causal bearing on the shape of any particular language, and thus on the diversity of human languages. While seemingly unprecedented in biology, language diachrony nonetheless recapitulates many general features of biological evolution. This may be summed up in the following observations:

- Today's micro-variation within the species or language is, at least potentially, tomorrow's macro-variation across both species or languages.
- Conversely, today's starkly diverse extant species or languages, genera or sub-families, and phyla or language families can be traced back to earlier, more humble variation at lower taxonomic levels (sub-species or dialect).
- Consequently, gradual step-by-step *micro-variation* can yield over time stark and seemingly unbridgeable gaps of *macro-variation* among extant species or languages.
- The process of change itself, the invisible teleological hand that guides the ever-shifting but still roughly isomorphic matching of structures to functions, is driven by adaptive selection (functional pressures).
- The overlaying of adaptively-driven changes in temporal order can lead, over time, to considerable restructuring and arbitrary structure-function pairing, thus to seemingly non-adaptive *relic features* (excess structure, spandrels).

Above all, both language diachrony and biological evolution abide by the following principles of developmental control:

- Graduality of change
- Adaptive motivation

- Terminal addition (of new structures to older ones)
- Local causation (but global consequences).

Since we have virtually no direct fossil evidence of prior stages of language evolution, the topic remains difficult and controversial. Yet true understanding of human language will not be possible without an evolutionary account. To some extent, and with a great measure of caution, the two well-documented developmental domains that are accessible to us, language diachrony and child language development, could furnish us with useful hints about possible evolutionary scenarios; as could, to some extent, the study of second language acquisition and pidginization (Givón 1979, 2002, 2005; Bickerton 1981, 1990; Botha 2005; Heine and Kuteva 2007).

2.6.2 The adaptive ecology of human communication

The rise of the two structural (symbolic) codes unique to human communication—phonology and grammar—is but the *adaptive response* to three more profound changes in the ecology of human communication. These changes constituted the *adaptive motivation* for the rise of human language as we know it, and are in turn themselves motivated by various facets of human cultural evolution.[3]

(a) Spatio-temporal displacement of reference

Both early childhood communication and pre-human—including primate—communication are heavily weighted toward here-and-now, you-and-I, and this-or-that referents that are perceptually accessible in the *immediate speech situation*. When all referents are equally accessible to all participants in the shared speech situation, the lexical coding of the *type* of referent is superfluous. Mere pointing (deixis)—orienting the interlocutor to achieve of *joint attention*—will suffice.

Mature human communication is, in contrast, heavily tilted toward spatio-temporally displaced referents, be they individuals, objects, state or events. This is reflected first in the lopsided use-frequencies of displaced reference. But it is also reflected, in turn, in the fact that much of our grammatical machinery is dedicated to communicating about displaced referents, states, and events (Givón 2001*a*).

Referents in the shared immediate speech situation are mentally represented in the working memory/attention system. Such representation shifts—with motion and attention—from one moment to the next, and is thus temporally unstable. In contrast, displaced referents are more likely to be representations in episodic

[3] For lack of space I will not be able to discuss here human cultural evolution. Many causal scenarios for language evolution have been suggested, most of them reduction to a single adaptive factor (descent from the trees, bipedism, widening foraging range, omnivorous feeding, tool-making, social grooming, increased group size, big-game hunting, big-game scavenging, laryngeal retraction, etc.). A more likely scenario was probably complex and interactive (co-evolutionary). For a more balanced discussion by non-linguists, see de Waal (2001), Geary (2005), Cheney and Seyfarth (2007).

memory, as either memories of past experience or future projections, plans, or imaginations. Compared to working memory, episodic memory is a much more stable mental representation. And this *temporal stability* may have contributed toward the objectivization of verbally-coded referents, including mental predicates.

The rise of the human lexical code—phonology—may now be understood as an adaptation designed to accommodate the shift to displaced reference in human communication. When the adaptively-relevant topics of communication became, increasingly, the spatially-displaced past experiences or future plans of some individual rather than of everybody present on the scene, pointing ceased to be a viable tool of referent identification.

(b) Declarative speech acts

Spontaneous pre-human communication is confined almost exclusively to manipulative speech acts (Tomasello and Call 1997; Savage-Rumbaugh et al. 1993; Pepperberg 1999; Cheyney and Seyfarth 2007), a tendency also observed in early childhood communication (Carter 1974; Bates et al. 1975, 1979; Givón 2008). In striking contrast, mature human discourse is tilted heavily, at the use-frequency level, toward declarative speech acts (Givón 1995, ch. 2), and the bulk of the grammatical machinery of human language is invested in coding declarative speech acts (Givón 2001*a*).

The emergence of declarative speech acts may have enhanced the liberation of epistemic mental predicates from their erstwhile subordination to deontic predicates. And the separate and more explicit representation of epistemic mental states ("think", "know", "see", etc.) may have, in turn, contributed toward heightened consciousness of *mental framing* operators, first those referring to one's own mental states, and then, by extension, those referring to the mental states of others.

The emergence of declarative communication also points toward the increasing adaptive relevance of displaced reference. Manipulative speech acts are confined to here-and-now, you-and-I, the immediate speech situation, that is, primarily to what is represented in working memory and current focal attention. Declarative and interrogative speech acts, on the other hand, are utterly superfluous when the referents are equally available to both interlocutors here-and-now. Why bother to tell the other guy if he already knows what you know? Why bother to ask them if you already know what they know?

It is the emergence of displaced reference as the more prevalent topic of communication that endows declarative (and interrogative) speech acts with their adaptive motivation: They are designed to carry the load of reporting (and querying) about inaccessible referents and past or future events that are not available to all interlocutors. Displaced reference creates an *informational imbalance* in the erstwhile intimate social unit, and declarative/interrogative speech acts are the adaptive response to such an imbalance.

(c) Multi-propositional discourse

Both early childhood and primate communication are overwhelmingly mono-propositional (Tomasello and Call 1997; Savage-Rumbaugh et al. 1983; Cheyney and Seyfarth 2007; Bloom 1973; Carter 1974; Scollon 1976; Bates et al. 1975, 1979; Givón 2008). In contrast, mature human communication is, at the use-frequency level, overwhelmingly multi-propositional. This is also reflected in the fact that the bulk of the machinery of grammar is invested in coding multi-propositional, cross-clausal coherence (Givón 2001*a*).

As noted above, grammar codes, primarily, the mental representation of the interlocutor's ever-shifting epistemic and deontic states during communication. The high automaticity of grammar may mean, among other things, that the evolution of grammatical communication was motivated, at least in part, by the strong adaptive pressure of having to deal with a high frequency of *perspective shifting*; perhaps an order of magnitude higher than what pre-human social species had to deal with.

One may view the rise of multi-propositional discourse as but the next step in the rise of declarative communication. As the volume of adaptively-relevant information about displaced reference became greater, the faster, more streamlined processing of such voluminous information became more adaptively pressing, especially in terms of the constant perspective-shifting involved in the processing of larger stretches of coherent discourse. The rise of grammar may be viewed as an adaptive response to the need to process this explosion of declarative multi-propositional information.

2.7 Typological diversity and language universals

To the naked eye, both the linguist's and the lay person's, the diversity of human languages seems immense and unconstrained, at first glance defying any attempt to posit meaningful universals. The argument whether features of human language are universal (and thus, by implication, motivated, genetically transmitted, biologically evolved and *innate*) or idiosyncratic (and thus, by implication, arbitrary, non-biological and culturally transmitted) harkens back to both Plato (*Cratylus*, *Meno*, *Phaedo*) and Aristotle (*De Interpretatione*, *The Categories*).

Many fine linguists, especially those who followed the structuralist dogma of arbitrariness (F. de Saussure, L. Bloomfield), expressed strong doubts about language universals, believing in unconstrained cross-language variation. Others, like

Chomsky, have militated for an extreme version of universality and innateness, by extracting from the vast and varied phenomenology of language features a few sufficiently abstract ones that are then said to be shared by all human languages. This gambit echoes both Plato's *eidon* (essence) and Saussure's *langue* (underlying system).

A more balanced empirical approach to the problem, perhaps best exemplified in the works of Joseph Greenberg (cf. Greenberg 1966*b*), adopts a middle-ground biological perspective, whereby both variation and universals are acknowledged. Indeed, the two are closely intertwined, and the balance between them is mediated by developmental processes. Thus, specific features of both phonology and syntax may vary considerably across languages, and the aggregation of such variation may lead to a, seemingly, staggering cross-language diversity. But within each functional-structural domain, the variation is severely constrained—say, five to seven major types of structures that usually code the same communicative function. And the constraints on variation are mediated by *general adaptive principles*, which in turn manifest themselves through the three developmental trends—language evolution, child language development, and diachronic change. This is a profoundly biological perspective, illuminating the fundamental unity of the disciplines that study life. As in evolution, then, language universals are not a set of concrete traits found in all languages but rather a set of general principles that control development and thus motivate the genesis of concrete traits.

2.8 Innateness and language acquisition

The controversy about innateness and language acquisition—whether the seemingly universal features of language are innate or are acquired from experience and input—harkens back to antiquity. In epistemology, a protracted debate took place between extreme *rationalists* (following Plato's lead in the *Meno* and *Phaedo* dialogs) and extreme *empiricists* (following Aristotle's *De Interpretatione*). The argument was resolved by Kant (*Critique of Pure Reason*), showing that both innate knowledge and experience are necessary ingredients of learning. Put another way, an organism never starts from scratch, it always has some pre-set mental parameters. That is, after all, what evolution is all about. But the pre-set parameters interact with individual experience, and ultimately the incrementation of knowledge is the product of such interaction. In linguistics, the argument was resurrected in the discussion of child language acquisition, with Bloomfield (1933) and Skinner (1957) taking the extreme empiricist position, and Chomsky (1959; 1965, ch. 1) arguing for an equally extreme rationalism. An empirical study of child language resolves the argument, again, toward the interactive middle.

In biology, where it all started, innateness is a mundane fact of life, the mere concession that evolution has indeed occurred, and that a species need not start its adaptive journey from the amoeba every new generation. For the crux of evolution is the genetic encoding of the cumulative adaptive experience of prior generations, so that it may be preserved and transmitted to subsequent generations. Without innateness, there is no inheritance of evolved adaptive evolved traits. But without individual behavioral experimentation, behavioral opportunism, and learning from experience, there is no adaptive selection, thus no evolution.

2.9 Language and the brain

The 1970s saw the emergence of an important but ultimately too simple account of the neurology of language, the so-called Geschwind (1970) model which located grammar in one left-cortical area (the pre-frontal Broca's area) and meaning in another (the medio-temporal Wernike's area). More recent work suggests that both meaning (Posner and Pavese 1997; Abdulaev and Posner 1997; Caramazza and Mahon 2006; Tucker 1992) and grammar are multi-modal distributive networks spanning many cortical and sub-cortical loci. In that, they follow the pattern of other higher cognitive networks, such as vision (Ungerleier and Mishkin 1982; Kaas 1989) and attention (Schneider and Chein 2003; Posner and Fan 2004).

While the neurology of grammar is yet to be fully resolved, the following connectivity is suggested by its functional interaction (Givón 2005, ch. 4):

(16) Grammar as a multi-modular distributive network:

Functional modules	Putative brain loci
(i) semantic memory	L-pre-frontal dorsal; sensory/affective sites
(ii) propositional semantics	L-medio-temporal (Wernike's area?)
(iii) episodic memory	hippocampus and amygdala
(iv) working-memory buffer(s)	R-posterior-parietal cortex (?)
(v) the attentional system	widely distributed
(vi) rhythmic-hierarchic processor	L-pre-frontal cortex (Broca's area)
(vii) grammatical morphology	L-pre-frontal cortex (?) (Broca's area?)
(viii) "theories of mind"	R-dorso-lateral-pre-frontal cortex (?)
(ix) automaticity processors	cerebellum and basal ganglia

2.10 The evolution of grammar

Grammar is at the very heart of the incredibly efficient, highly automated human communication system. It is a rapid processor for encoding the interlocutor's rapidly shifting deontic and epistemic states during ongoing communication. For the evolution of grammar to be neurologically plausible, it must be understood as the extension and elaboration of well-documented precursors. The following is part of what must be accounted for, focusing on the representation of other minds ("theory of mind"):

(a) Conscious, intentional organisms have a—probably automatic—co-activation relation between their motor system and their intention-encoding system.
(b) The work of Rizzolatti et al. (1996*a*, 1996*b*, 2000), Milner and Goodale (1995), and Goodale (2000), taken together, indicate automatic mutual co-activation of the motor and visual systems.
(c) In particular, the mirror-neuron literature suggests that viewing the actions of another person automatically activates one's own motor system. That, given (a) above, would presumably also activate, automatically, one's intentional system.
(d) There are a lot of indications that monkeys, let alone chimps and bonobos, already view their conspecifics as intentional beings (Cheney and Seyfarth 2007); that is, as beings whose actions are co-activated by or with their intentions.
(e) All this, taken together, suggests a natural mechanism whereby viewing the interlocutor's actions will also activate a mental model, in the viewer's mind, of the interlocutor's intentions.
(f) If it is true, as some have suggested, that grammar in its evolution piggy-backed on both the visual system (Givón 1995, 2002) and motor system (Greenfield 1991), then grammar as an instrument of rapid coding of the interlocutor's intentional states during ongoing communication would now have a natural evolutionary pathway.
(g) Both evolutionary neurology (Mesulam 2000) and facts of language development (Bates et al. 1975; Givón 2008) and diachrony (Givón 1979; Heine et al. 1991) suggest that epistemic meaning earlier are attached to deontic (intentional) meanings, and only later "liberate" themselves.[4]

[4] Premack and Woodruff (1978) suggest that the mental representation of epistemic states was a later addition to the representation of deontic states. However, the intentional "I'd like to eat the apple" presupposes the two epistemic states, (i) the factual current state "I haven't yet eaten the apple", and (ii) the intended future state "I will eat the apple". As in the diachronic development, where epistemic senses are a later development out of deontic ones, evolution simply liberates the deontic from its dependence on the deontic. The relation between the two is thus a one-way implication: DEONT $\supset$ EPIST. In diachronic terms, this is "liberation" or "bleaching". Likewise, in the child's acquisition of propositional modalities, deontic modalities are acquired earlier than epistemic ones (Diessel 2005; Givón 2008), and epistemic usage is often "liberated" from earlier deontic usage.

(h) The addition of encoding the interlocutor's epistemic (belief) states would thus be a later development, one not attested in monkeys but perhaps attested in a rudimentary way in chimpanzees and bonobos (Cheyney and Seyfarth 2007). Such a "liberation" sequence is still recapitulated in both the ontogeny and diachrony of language.

Obviously, the precise neurological details of this scenario, and the connection between brain mechanism and syntactic constructions, are yet to be resolved.

2.11 Theoretical perspective and grammatical description

One may well ask what this, or any other, theoretical perspective has to do with describing the grammar of a particular language? The answer is that descriptive practices and methods are always anchored, whether consciously or not, in distinct theoretical presuppositions. The type of classificatory categories brought to bear on primary language data are highly theory-laden, as are the very assumptions of what constitutes relevant primary language-data to begin with, and what is the appropriate method for analyzing it. Thus, Chomsky's rationalist perspective on language and learning would place a heavy value on the introspected knowledge of the native speaker, and thus on clauses detached from their natural communicative context. An adaptive perspective, on the other hand, would note that our knowledge of grammar is mostly implicit and subconscious, and that it must therefore be gleaned, ultimately, from the study of natural communication.

In describing a language, the functionally- and typologically-oriented linguist takes it for granted that there are universal categories that are *most likely* to be found in most, if not all, natural languages. But he or she must also remain mindful of the possibility that such categories are not absolute, that some are more universal than others, that language diversity is not trivial, and that any language has a range of options to choose from in grammar-coding the same communicative tasks. What is more, the same language may avail itself of a range of *alternative structural means* for performing the same—or roughly similar—communicative function. And those alternative adaptive solutions often become prominent sequentially through the historical development of a grammar.

Likewise, in the few "lexical" signals of pre-human communication, such as mating or predator calls, the usage is always manipulative (deontic), never referential (epistemic; Cheney and Seyfarth 2007).

The theoretical perspective I have elucidated above would begin describing a language by looking at the *lexicon* and first at words that describe the relevant physical and social universe. It is most natural to begin with nouns—physically compact, time-stable, concrete entities, those whose meanings are likely to be apparent regardless of the specific event/state context in which they are used: body parts, kinship, and gender/age terms, features of the terrain, flora and fauna, artifacts, clothes and habitations, celestial and weather phenomena.[5]

Once enough nouns and their approximate meanings have been gathered, one has constructed, *prima facie*, a simple physical-cultural universe in which various types of events, states, and behaviors can now be situated. The lexicon of predicates—verbs and, if relevant, adjectives—can be now investigated by eliciting simple propositions (main, declarative, affirmative, active clauses) in which the nouns may partake as agents, recipients, patients, locations, instruments, etc. (semantic roles) in simple events or states.

The various types of events/states will eventually yield the corresponding types of *simple structures*, so-called "verb-frames" or "argument structure". And this in turn may yield a preliminary scheme of the grammatical coding of event participants as subjects, direct objects, or indirect objects (grammatical roles). Given that some verbs ("want", "order", "know") can take clausal complements, this stage of the investigation would also reveal, again in a preliminary fashion, one general type of complex construction, that which takes verbal complements.

At this early phase of the investigation, it is advantageous not to deal with larger, multi-propositional chunks of coherent discourse, although one must constantly keep those in mind as the ultimate context for using clauses in natural communication. The preliminary description of grammatical constructions at this early phase of the investigation is thus likely to be only partially accurate, in particular as far as communicative function is concerned. But this kind of partial accuracy is a necessary early step in all science, where one perforce starts in restricted domains and simplified, somewhat artificial sub-systems. This is the way of *controlled experiment* in science or *in vitro* studies in biology. One does this in language description knowing full well that the realistic context of natural communication is much more complex, and would require revision and expansion of the preliminary grammatical description obtained from the study of isolated clauses.[6]

The natural communicative context may be eased in gradually, beginning with the few non-simple (transformed) clause-types whose context is a little more transparent cross-culturally: negation, questions, commands, tense, modality. Thus, starting with the simple clause (17a) below, one may obtain (17b–g):

[5] A fringe benefit of this early stage of the investigation is that it supplies most of the evidence for resolving the sound system (phonology) of the language.

[6] Invoked or uninvoked, the communicative context is—much as Xenophon's God would have been—always there.

(17) *a*. Simple: Marvin gave the book to Marla.
b. Negative: Marvin **didn't** give the book to Marla.
c. Potential: Marvin **may** give the book to Marla.
d. Command: Give the book to Marla!
e. Request: **Would** you **please** give the book to Marla?
f. WH-question: **Who** gave the book to Marla?
What did Marvin give Marla?
g. Y/N-question: **Did** Marvin give the book to Marla?

One can then gradually build up the communicative context around the target proposition to see how its structure adjusts accordingly. Thus, in trying to elicit something like the contrast between indefinite and definite noun phrases in English, one may negotiate hypothetical communicative scenarios such as, for example:

(18) Eliciting a referring indefinite noun phrase:

"Suppose you met some woman yesterday, one I know nothing about, and you found her very attractive and you want to tell me about her, how would you go about beginning the telling?"

This may be then contrasted with:

(19) Eliciting a definite noun phrase:

"Suppose you already told me about meeting that woman, and now you want to talk about her again, how would you do it?"

There are many other strategies that may be used in constructing more and more natural communicative contexts. But the ultimate strategy is, still, to record natural conversations, narratives, procedural descriptions, etc., and then to analyze the deployment of grammatical constructions in the various communicative contexts. This is the penultimate test for understanding grammar, the one most available to linguists. The ultimate will be, some day, controlled experimental methods.

The functional and typological approach to grammar helps the descriptive endeavor at both ends of the communicative (semiotic) equation. At the functional end, it furnishes us with a universal inventory of *communicative functions* (functional domains), those that are most likely to be performed by grammatical constructions in natural languages. At the structural end, this approach furnishes us with an inventory of the types of *grammatical structures* that are most likely to code any particular communicative function. The underlying logic of *grammatical typology* is just that (Givón 1995, ch. 3): (i) first elucidate the universal inventory of grammar-coded communicative functions; and (ii) then elucidate the range of grammatical variants that can code each *functional domain*. However, both universal inventories still need to be explained, presumably by a full-fledged theory that takes account of—at the very least—the findings of communication, cognition, neuro-biology, and anthropology.

This page intentionally left blank

CHAPTER 3

THE CARTOGRAPHY OF SYNTACTIC STRUCTURES

GUGLIELMO CINQUE
LUIGI RIZZI

3.1 Introduction

Syntactic structures are complex objects, whose subtle properties have been highlighted and elucidated by half a century of formal syntactic studies, building on a much older tradition. Structures are interesting objects of their own, both in their internal constitution and in their interactions with various grammatical principles and processes. The cartography of syntactic structures is the line of research which addresses this topic: it is the attempt to draw maps as precise and detailed as possible of syntactic configurations. Broadly construed in this way, cartography is not an approach or a hypothesis, it is a research topic asking the question: what are the right structural maps for natural language syntax? Answers may differ, and very different maps may be, and have been, proposed, but the question as such inevitably arises as a legitimate and central question for syntactic theory. If it is a virtual truism that cartography can be construed as a topic and not as a framework, it is also the case that cartographic studies have often adopted certain methodological and heuristic guidelines, and also certain substantive hypotheses on the nature of syntactic structures, which form a coherent body of assumptions and a rather well-defined research direction; we will try to illustrate some ideas and results of this direction in the present chapter.

If structures have, in a sense, always been central in generative grammar, the idea of focusing on structural maps arose around the early 1990s, following a track parallel to and interacting with the Minimalist Program. Perhaps the main trigger was the explosion of functional heads identified and implied in syntactic analyses in the first ten years of the Principles and Parameters framework. One critical step was the full-fledged extension of X-bar theory to the functional elements of the clause (Chomsky 1986*b*) as a CP–IP–VP structure; and the observation that other configurations, e.g., nominal expressions, were amenable to a hierarchical structure with a lexical projection embedded within a functional structure (such as Abney's DP hypothesis, Abney 1987). These advances provided a natural format for the study of the structure of phrases and clauses as hierarchical sequences of the same building block, the fundamental X-bar schema (or, later, elementary applications of Merge); the lowest occurrence of the building block typically is the projection of a lexical category, e.g. a noun or a verb, and this element is typically completed by a series of building blocks headed by functional elements, providing more abstract semantic specifications to the descriptive content of the lexical head: tense, mood, aspect for the verb, definiteness, specificity, number for the noun, etc.

If the first step was the idea that clauses and phrases are formed by a lexical structure and a higher functional structure, both corresponding to elementary building blocks hierarchically organized, the second crucial step was the observation that the functional structure typically consists of more than one head. In fact, a Complementizer Phrase (CP) and an Inflectional Phrase (IP) zone were isolated from the outset, but it became clear very soon that the same kinds of evidence which supported the analysis of inflected verbs in terms of the distinction between I and V would lead to the splitting of I into more elementary components. The same logic led to a later splitting of the CP and DP zones into more articulated hierarchical sequences of functional projections.

The initial impulse for splitting the IP was provided by Pollock's seminal paper on verb movement in French and English (Pollock 1989, versions of which circulated already around the mid-1980s). Pollock showed that assuming a single I position did not provide enough space to account for the different positions that can be occupied by different morphological forms of the verb in French: infinitival verbs may remain in the VP, as in (1)a, or be moved to a higher position across lower adverbs like *complètement* (completely), as in (1)b; finite verbs move to an even higher position across negative *pas*, as in (1)c:

(1) a. ne X_1 pas X_2 complètement [$_{X3}$ comprendre] la théorie...
Neg not completely understand the theory
b. ne X_1 pas [$_{X2}$ comprendre] complètement X_3 la théorie...
c. Il ne [$_{X1}$ comprend] pas X_2 complètement X_3 la théorie

If I splits into at least two heads X_1 and X_2, Pollock argued, the three positions of (1) can be naturally accommodated by assuming optional movement of the infinitival verb from its VP-internal position X_3 to X_2, and obligatory verb movement of the finite verb to X_1. This analysis, also building on Emonds (1978), introduced a fundamental insight: adverbs basically don't move, except in the cases in which they are displaced for scope-discourse reasons, focalized, and the like; variations within a language and across languages of verb–adverb orders are due to verb movement in the inflectional space, a particular instance of head movement. This approach in fact united two lines of research which have become integral components of the cartographic studies: on the one hand, the analysis of the word order properties of verbs with respect to adverbials and arguments in terms of head movement, as mentioned above; on the other hand, the idea that inflectional morphology is done in the syntax and is the result of movement rules involving roots and affixes, an idea going back to the analysis of verb affixation in English in *Syntactic Structures* (Chomsky 1957). The Emonds–Pollock approach united the two trends by proposing that the verb could be attracted to different functional positions to pick up affixes and get properly inflected, thus changing its position with respect to adverbs and other elements, which made it possible to capture many important form-position correlations.

The question then arose of the proper labeling of X_1 and X_2. Belletti's (1990) proposal was that the higher functional projection of the clause is the one responsible for subject–verb agreement (AgrS in the traditional terminology), and the lower one expresses tense. This order AgrS–T is immediately reflected in the order of prefixes or particles in, for example, the Bantu languages; while in languages in which these properties are expressed by suffixes, i.e., the Romance languages, the order is the mirror image (see Italian *parla-v-ano*, root-T-AgrS, "(they) spoke"), as is to be expected under Baker's (1985) Mirror Principle: the verb moves to pick up the closest suffix, which therefore appears as the one immediately attached to the stem, etc.

The logic of this argumentation, combining the syntactic make-up of inflectional morphology via head-movement and the study of the order of arguments and adjuncts with respect to different verbal forms, quickly led to a finer splitting of the inflectional space into a sequence of functional heads expressing properties of mood and modality, tense, aspect, voice. For a few years, around the late 1980s, this methodology led to the discovery and postulation of a variety of functional heads driven by the analytic necessities of particular morphosyntactic problems, a trend which sometimes gave the impression that the process would lead to an ever-increasing complexity of the syntactic representations. How rich could be the "right" functional structure of clauses and phrases? One of the driving ideas of the cartographic projects was precisely to complement this trend of bottom-up, problem-related discovery with a more top-down, global perspective, trying to make a rough estimate of the upper limit of the structural complexity. Instrumental to this endeavor was the working assumption that each morphosyntactic feature would correspond to an independent syntactic head with a specific slot in the functional hierarchy (cf. also Kayne 2005*a*: 15). Much of the cartographic work has consisted in the attempt, in various forms, to use this working hypothesis as a heuristic guideline, thus spelling out empirical arguments supporting or disconfirming its validity across languages.

3.2 Methodology and evidence

In the first half of the last century, in part as a reaction to what was then felt as an unwarranted application of European grammatical categories and constructions to non-European languages, the common wisdom in American structuralism (epitomized in Joos 1957: 96) was that "languages could differ from each other without limit and in unpredictable ways" so that each language should be studied "without any preexistent scheme of what a language must be". The rejection of these assumptions, which are still adopted today by many functionalists,[1] was implicit throughout the history of generative grammar,[2] and is made explicit in Chomsky's (2001*b*) "Uniformity Principle" ("In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances."). The cartographic approach follows this idea in assuming that all languages share the same principles of phrase and clause composition and the same functional make-up of the clause and its phrases.[3]

[1] See, for example, LaPolla and Poa (2002: 2): "Each language is a unique set of language-specific conventions, and so each language should be described in its own terms."; or Haspelmath (2007; this volume): "Descriptive linguists still have no choice but to adopt the Boasian approach of positing special language-particular categories for each language. Theorists often resist it, but the crosslinguistic evidence is not converging on a smallish set of possibly innate categories. On the contrary, almost every newly described language presents us with some 'crazy' new category that hardly fits existing taxonomies."

[2] See, for example, Koopman and Sportiche (1991: 218f.): "[W]e suppose that the null assumption concerning language variation is that it does not exist."

[3] This is not to say that it is always easy to establish precise correspondences between the functional categories overtly displayed by different languages. Caution must be exercised, but there is

More precisely the cartographic approach assumes, as the evidence of the last several years seems to indicate, that the distinct hierarchies of functional projections dominating VP, NP, AP, PP, IP, etc. may be universal in the type of heads and specifiers that they involve, in their number, and in their relative order, even if languages differ in the types of movements that they admit or in the extent to which they overtly realize each head and specifier (Rizzi 1997; Cinque 1999; 2002: 3f.). This is the strongest position one could take, one which implies that if some language provides evidence for the existence of a particular functional head (and projection), then that head (and projection) must be present in every other language, whether the language offers overt evidence for it or not (cf Kayne 2005a: 12; Cinque 2006*a*: 4).[4]

A weaker position would consist in assuming that languages may differ in the type or number of functional projections they select from a universal inventory, or in their order.[5] Although the choice between these two positions will ultimately be decided by the nature of things, methodologically it would be wrong, it seems, to adopt the weaker position as a first working hypothesis. That would only make us less demanding with respect to the facts and could lead us to miss more subtle evidence supporting the stronger position, a risk not present under the other option (Cinque 2002: 4).

The question whether such universal hierarchies of functional projections are primitive objects of UG (Universal Grammar), or can be derived from interface or more general external conditions, is important, but fundamentally orthogonal to the prior task of drawing their precise map, and perhaps not easily determinable in the present state of our knowledge.

The evidence brought to bear in the literature on the mapping of universal hierarchies of functional projections comes from a variety of sources. An early source for postulating (abstract) functional projections was the existence of certain systematic word order differences among languages, such as Pollock's (1989) classical argument for positing a non-lexical head higher than VP and lower than I (or T), to which finite verbs raise in French (but not in English), along the lines discussed in the introductory section.

Another important source of evidence is the relative order of the functional morphemes overtly realized in the languages of the world (to the extent

no a priori reason to rule out the possibility that such correspondences can ultimately be established. In fact, this has turned out to be possible in a number of cases through in-depth investigation. See, for example, the case of French *peu* and English *bit* (rather than *little*) discussed in Kayne (2005a, §4.2).

[4] The literature offers a number of cases supporting this general hypothesis. See, for example, the discovery of more subtle evidence for the presence of a DP projection in languages like Serbo-Croatian, Russian, and Japanese, which lack overt determiners (Progovac 1998; Pereltsvaig 2007; Furuya 2008); or the indirect evidence discussed in Kayne (2003: 219) and Cinque (2006*b*) for the presence of numeral classifiers in languages like English and Italian, which are traditionally taken not to be numeral classifier languages.

[5] This is the position taken, for example, by Thráinsson (1996) and Bobaljik and Thráinsson (1998), among others. See also Fukui (1995).

that one can establish reasonable correspondences among the functional morphemes of different languages). Though languages differ as to what functional categories they overtly realize, the partial orders displayed by different languages seem to fit in a unique macro-hierarchy despite occasional inconsistencies, which have proved (and hopefully will prove, as our knowledge progresses) solvable.

Preliminary inquiries on the functional hierarchies of the left periphery of the clause (Rizzi 1997; 2001; 2004*a*; 2004*b*; Benincà 2001; 2006; Benincà and Poletto 2004; Bocci 2004; Benincà and Munaro to appear; Cruschina 2006; 2008; Frascarelli and Hinterhölzl 2007; Frascarelli and Puglielli to appear, among others), of the core functional structure of the clause (Cinque 1999; 2006; Shlonsky 1997; 2000; Sigurðsson 2000; Cardinaletti 2004; Schweikert 2005; Bianchi 2006; and, for its relevance for computational linguistics, Chesi 2005), of the DP (Cinque 1994; Scott 2002; Brugè 2002; Giusti 2002; Nicolis 2008; Svenonius 2008*a*), and of PPs (see the contributions in Asbury, Dotlačil, Gehrke, Nouwen 2008, and Cinque and Rizzi to appear) have largely confirmed the working hypothesis that there may be a universal functional design for the clause and its major phrases holding across languages.[6]

Of course, to determine the relative order of functional morphemes one has to have an idea of what the classes of such elements are as opposed to the lexical ones (see section 3.3 below for discussion), and this task often requires "regularizing" the orders found across languages, as they can be obscured to various degrees by various types of syntactic movements. So, for example, the relative order of functional morphemes that appear to the right of a certain lexical category, as suffixes or free morphemes, is most often (though by no means always) the mirror image of the same functional morphemes that appear to the left of the same lexical category in other languages, arguably a consequence of the lexical category moving across the functional morphemes in the former types of languages (see Baker 1985 for the original formulation of the Mirror Principle and, for recent discussion, see Cinque 2009).

[6] Some authors have argued that this particular assumption of the cartographic approach is incorrect because it rests on transitivity (if A > B and B > C, then A > C), which appears to fail in certain cases (see Bobaljik 1999; Nilsen 2003, and also Zwart 2006). Caution, however, is in order given the otherwise general validity of transitivity, and the possibility that some account exists which renders these cases irrelevant for transivity issues. See in fact Cinque (2004, footnotes 22 and 43 for evidence to this effect). Van Craenenbroeck's (2006) analogous argument from an apparent transitivity failure in the left periphery also ignores the possibility that the complementizer may occupy more than one position, thus rendering his case irrelevant to the transitivity issue. That an element like *that* may appear more than once and in different positions in the left periphery of a clause is straightforwardly shown by many cases of multiple occurrences of *that*, e.g., in Brazilian Portuguese, Gascon, and Piedmontese structures with orders like *I think that JOHN that you should meet*, with the first *that* functioning as declarative force marker, and the second as a focus marker (see Mioto 1998, Poletto 2000: 148–50 for relevant discussion).

Analogously, as noted in Carlson (1983: 73), one of the earliest and most enlightening discussions of functional categories in the generative tradition, the Latin coordinating enclitic conjunction *-que* exemplified in (2), is not interpreted as conjoining with a like constituent that precedes it (i.e., the unit [*ob eās*]) but the entire higher unit [*ob eās rēs*] (as in English). This again can be "regularized" if the movements that created (2) (from ... *ob eās rēs -que*) are undone.

(2) *ob* *eās-***que** *rēs*
because.of these-**and** things
'and because of these things'

These are two of the many cases where care must be taken to render things comparable and to expose the deeper regularities that underlie the functional make-up of the clause and its phrases.

3.3 Inventory of functional categories

A guiding idea of much current cartographic work is that the inventory of functional elements (heads or specifiers of functional projections) is much larger than is generally thought. In all grammatical traditions it is customary to make, in one way or another, some distinction between lexical categories (like Nouns and Verbs: see Baker 2003) and functional, or grammatical, ones (like Determiners and Complementizers). If we take membership in an open vs. closed class of items as a diagnostic distinguishing lexical from functional elements, then the candidates for the functional lexicon of languages become very numerous. Not only are Determiners and Complementizers functional but also conjunctions, (functional) adpositions like *of, for, from, at, to, with* (as well as spatial adpositions—see Cinque and Rizzi to appear, and references cited there), mood, modal, tense, aspect, polarity, and voice morphemes,[7] auxiliaries, copulas, and other verbs lacking a clear argument structure, (strong, weak, and clitic) pronouns, demonstratives, quantifiers, numerals (see Kayne 2005a: 13), classifiers, number (plural, dual, etc.) morphemes, gender or class morphemes, diminutive/augmentative morphemes, degree words, indefinite/*wh*-words, case morphemes, focusing adverbs (like "only" and "also"), comparative and superlative morphemes, and many many more (see Kayne 2005a, section 2.1). To judge from Heine and Kuteva's (2002) four hundred or so independent grammaticalization targets, the number of functional elements must at least be of that order of magnitude. It is in fact easier to consider which elements are lexical (belong to

[7] Whether bound or free. On the (functional) syntactic import of bound morphemes, see the recent discussion in Kayne (2005a: 11f.).

an open class). Nouns in all languages appear to be an open class, perhaps the only genuinely open class, as the considerations that follow may indicate. The situation is certainly far less clear for adjectives, adverbs, and verbs (which are often taken to be lexical, open classes). In many languages, adjectives constitute a closed, often quite small, class of elements. This is especially evident in those languages, like Yoruba (see Cinque 2006*a*: 5 and references cited there), whose adjectives cannot be used predicatively. In such languages the attributive-only adjectives form a closed (generally small) class, a clear sign of their functional status. For discussion and exemplification, see Dixon (1982; 2004), Cinque (2006*a*: 4f., to appear). The fact that they appear to form an open class in other languages may be due to the existence of a parallel predicative class of adjectives (which enlarges the set of adnominal adjectives by adding a reduced relative clause source), as well as to possible productive morphological derivations of adjectives from nouns or verbs (e.g., *-al, -ous, -ed*, etc. in English).

A similar situation is encountered with adverbs, which also clearly constitute a closed class of elements in some languages (see Dixon 1982: 40; Schachter 1985: 21ff.; Stutzman 1997: 75; Cinque 1999: 213, fn. 79; 2006: 9, fn. 22, and references cited there). Furthermore, the fact that they are coded as rigidly ordered affixes in certain languages while they are coded as independent words in others (also in a fixed order) may suggest that generation in head or specifier position of a dedicated functional projection is an option left open to languages by UG.

If Hale and Keyser's (1993) idea that most transitive and intransitive verbs are not primitive but result from the incorporation of a noun into a limited class of light/general purpose verbs ("do", "give", "take", "put", "hit", etc.), then even the class of *primitive* verbs may turn out to be closed and relatively small. This seems to be confirmed by the fact that some languages typically fail to incorporate the noun into the light verb so that most "verbal meanings" are expressed as V + N periphrases. This is, for example, the case of Persian.[8] The typological literature also reports the case of a number of languages from Australia and New Guinea with closed classes of main verbs (see Dixon 1982: 225; Pawley 2006).[9]

[8] "Most verbal constructions in Persian are formed using a light verb such as *kardan* ('do', 'make'), *dâdan* ('give'), *zadan* ('hit', 'strike'). The number of verbs that can be used as light verbs is limited, but these constructions are extremely productive in Persian." (Megerdoomian, n.d.). Also see Karimi-Doostan (1997).

[9] Interestingly, the literature on agrammatism reports the fact that even main verbs are impaired. See Miceli et al. (1984) (thanks to Franco Denes for pointing out this article to us; there are also cases of selective impairment of the nominal system with verbs relatively spared (Caramazza and Shapiro 2004), but these are much rarer than cases of selective V impairment). If main verbs are the morphological merge of a noun plus one of a closed class of "grammatical" verbs, their conclusion that "agrammatism is a heterogeneous disorder that implicates damage of both lexical and syntactic mechanisms" (p.220) may have to be reassessed, and perhaps reduced to a disorder of (different types of) purely grammatical mechanisms.

3.4 Comparative syntax and typology

Crucial to the cartographic approach is the evidence coming from comparative and, more broadly, typological studies. These alone may help single out the variety (and the limits) of the functional lexicon of UG. In-depth studies of a single or of few languages, however deep they may be, fall short of revealing the actual richness of the functional/grammatical structure of UG owing to the often silent character of a certain functional category in a certain language (see Kayne 2005*a*; 2006). More importantly still, as noted, comparison of many different languages may provide evidence for determining the precise relative order of the different functional projections by combining the partial orders overtly manifested by different languages into what, in principle, should be a unique consistent order/hierarchy imposed by UG. This presupposes that the order of functional projections is fixed within one language and, more crucially, across languages—hardly an obvious assumption.

Comparative evidence is also crucial in exposing how certain ordering properties are strictly impossible across languages. Even in cases in which variation is permitted by UG, it is never the case that "anything goes". There are precise limits to the observed cross-linguistic variation, a fact which calls for a principled explanation. Consider, for example, the order of demonstratives, numerals, and adjectives with respect to the N (Greenberg's 1963 Universal 20). Even if variation in their relative ordering is extensive, of the twenty-four mathematically possible orders of the four elements, only thirteen are clearly attested in the languages of the world; apparently only those orders which are obtainable from a unique base order (Dem Num A N) by moving the N (or NP) leftward to higher functional positions in one of the ways independently admitted by the syntax of natural languages (see Cinque 2005 for discussion).

3.5 Cartography and Minimalism

The cartographic projects have been developed roughly at the same time as the rise and development of the Minimalist Program (Chomsky 1995 and much subsequent work). There is, at first sight, an inherent tension between the complexity of the cartographic representations and the simplicity of the generative devices that minimalist syntax assumes, somehow reflected in the structural poverty of the representations typically found in the minimalist literature. We believe that there is no contradiction between these two directions of research, and the tension, where real, is the sign of a fruitful division of labor. Minimalism focuses on the elementary mechanisms which are involved in syntactic computations, and claims that they can be reduced to extremely simple combinatorial operations, ultimately external and internal Merge, completed by some kind of search operation (Chomsky's Agree) to identify the candidates of Merge. An impoverished computational mechanism does not imply the generation of an impoverished structure: a very simple recursive operation can give rise to a very rich and complex structure as a function of the inventory of elements it operates on, and, first and foremost, of its very recursive nature. The very simplified structural representations often assumed in the minimalist literature, expressed by the C-T-v-V system, are sometimes taken literally as substantive hypotheses on the nature of clausal configurations, but the structure of the arguments rarely implies a literal interpretation, and often is compatible with an interpretation of C-T-v-V as a shorthand for more complex cartographic structures (a fact explicitly acknowledged, for example, in Chomsky 2001*b*, fn. 8), with C, T, and v taken as "abbreviations" standing for complex zones of the functional structure. The division of labor here is that Minimalism focuses on the generating devices, and cartography focuses on the fine details of the generated structures, two research topics which can be pursued in parallel in a fully consistent manner and along lines which can fruitfully interact (see Cinque 1999, section 6.2; Rizzi, 2004*a*, introduction, and Belletti 2009, introduction, for relevant discussion).

In fact, cartographic studies are based on general guidelines which are at the heart of the Minimalist Program. Minimalism has introduced a principled typology of UG principles, which are traced back to only two kinds of broad categories: principles dictated by the needs of the interface systems, determining the proper legibility and usability of the interface representations, and economy/locality principles, constraining the functioning of the computing machine.

The first class includes principles determining the mapping of a hierarchical structure onto a linear sequence expressible by the human articulatory system, such as Kayne's (1994) Linear Correspondence Axiom and its variants; and principles ensuring the expressibility of properties required by the human conceptual-intentional systems and by the needs of an efficient communication: properties of argument structure, referential dependencies, scope, and informational packaging in discourse and dialogue. All these aspects play a critical role in cartographic studies. Much work on the reordering of elements generating superficial exceptions to the hierarchical order crucially makes extensive use of remnant movement (e.g., Cinque 1999; Koopman and Szabolcsi 2000, but also much work on the left peripheral positions of *wh*-operators in Romance languages and dialects by Munaro, Obenauer, Poletto, Pollock), a direct offspring of the antisymmetric approach. Work on the cartography of the verbal system (Ramchand 2008) and of prepositions (Svenonius 2008*b* and the contributions collected in Cinque and Rizzi to appear) investigates the syntactic correlates of argument structure in structural approaches to the lexicon–syntax interface inspired by Hale and Keyser's (1993) perspective. Much work on the fine structure of the left periphery investigates the syntax of dedicated scope-discourse positions in various languages: Romance (Rizzi 1997; Benincà and Poletto 2004, and many other contributions in Rizzi 2004*a* and, on Romance dialects, Manzini and Savoia 2005), Germanic (Grewendorf 2002; Haegeman 2006), West African languages (Aboh 2004; 2007), Creole languages (Durrleman 2007), East Asian languages (Endo 2007; Tsai 2007).

The study of locality/economy is also central to the cartographic endeavor, in that the positional articulation uncovered by cartographic studies offers a sound basis for establishing a principled typology of positions which is required by the analysis of intervention locality: within the Relativized Minimality tradition (Rizzi 1990), an intervener of "the same kind" as the target of movement blocks a movement chain; the typology of positions cannot be established in the traditional terms of the A/A′ distinction (too coarse) nor in terms of a featural identity between the target and the intervener (too selective), and seems to require a feature-driven typology of an intermediate level of granularity, which can be directly related to the cartographic structures (Rizzi 2004; Starke 2001; Grillo 2008).

One point in which cartographic studies seem to us to fruitfully implement general simplicity guidelines which are peculiar to Minimalism is the study of the elements of syntactic computations. One useful heuristic principle which has guided much cartographic work is the maxim "one (morphosyntactic) property—one feature—one head". This guideline does not exclude the possibility that featurally complex heads may arise in syntax, but they cannot be "atoms" of the syntactic computations; they can only arise through derivational procedures, namely head movement, which may create a complex conglomerate of features by moving featurally simple heads into other heads (it does not matter here whether head movement literally extracts a head from its projection or is a kind of phrasal movement "in disguise"). It is this kind of intuition which guided the "unpacking" of the Infl node of early P&P analyses into its elementary components. Of course, a single surface position may express both the lexical content, tense, mood and subject agreement (as Italian present subjunctive *part-a-no* "that they leave"), but this is done through movement of the verbal head picking up the various elementary specifications. Similar considerations hold for the unpacking of the C node, of the determiner system, etc.

The basic intuition that cartographic studies try to validate empirically is that natural language design opts for local simplicity whenever possible: each syntactic head has a simple featural specification and can enter into few simple relations with its associates. Preservation of local simplicity is the effect massively produced by the pervasive presence of movement in natural language syntax. Consider for instance A′ movement chains, configurations which transparently arise to associate two kinds of interpretive properties to certain expressions. Thus, the expression *this* *book* must be interpreted as the thematic argument of the verb *read*, and as the topic of the structure in (3):

(3) This book, I will read ___ tomorrow

Natural languages express this state of affairs by having the element occur twice, once in the thematic position and once in the left peripheral position dedicated to topicality. The assignment of argumental thematic properties is, uncontroversially, a matter of head-dependent relation: the verb assigns a certain thematic role to its immediate dependent. What about a scope-discourse property like topicality? The line pursued by cartographic studies is that scope-discourse properties are assigned to elements in a configurationally uniform way, *mutatis mutandis*: there is a dedicated head, Top, normally occurring in the left periphery of the clause, which activates the interpretive instruction "my specifier is to be interpreted as the topic, and my complement as the comment". Under the copy theory of traces the full representation of (3) is

(4) This book [Top [I will read <this book> tomorrow]]
Topic Comment

with the silent copy in object position notated within angled brackets (on traces as silent copies see Chomsky 1995; Sportiche 2007, and others). Each head expresses a single property, we do not have complex heads simultaneously assigning to their dependents the complex of properties "patient of the verb *and* topic of the clause": natural languages opt for local simplicity, simple featural specifications on heads and local attribution of simple interpretive properties, even though the price to pay is a certain increase of global complexity, a richer functional structure and the multiple occurrence (or "movement") of an element in distinct structural positions. Similar considerations hold for other types of A′ constructions such as focus, questions, relatives, exclamatives, comparatives, etc.

A brief comment on representations like (4). The postulation of a Top head is immediately supported by the fact that in many languages a Top marker is in fact morphologically realized, i.e., Gungbe *yà* (Aboh 2004; 2007), Japanese *wa* (for a particular kind of topic), etc. A partial analogy can be drawn between such left-peripheral markers for scope-discourse semantic properties (topic, focus, Q, etc.) and inherent case for argumental properties (instrumental, locative, benefactive, ...): both morphosyntactic entities mark certain interpretive properties of one or the other kind, and both may superficially vary across languages in that they may or may not have a morphophonological realization.

This conception of A′ configurations implements in a very straightforward way the minimalist guideline according to which movement is a device to express an interface effect and, more generally, that linguistic computations are driven by the satisfaction of certain expressive needs of the interface systems (Fox 2000; Reinhart 2006). Among the advantages of this way of looking at things is the fact that A′

movement conforms to the general fact that movement is formally triggered by the featural constitution of a c-commanding head. More importantly, this conception makes possible a very transparent approach to the interface between syntax and semantics-pragmatics: peripheral functional heads can be seen as overt "flags" carrying very transparent instructions to the interface systems on how their immediate dependents are to be interpreted: (on the PF interface see Bocci 2009).

An objection that is sometimes raised against this view is that it seems to threaten the thesis of the autonomy of syntax. Granting the historical importance of the autonomy thesis in the process of properly structuring a rigorous and well-defined theory of syntax, we fail to see the force of this objection. First of all, we do not see why this conception should be perceived as more of a threat to the autonomy of syntax than the Theta Criterion, or the Projection Principle, or the theta-related character of inherent case assignment, or any other principle aiming at illustrating the transparency (ultimately, the simplicity) of the mapping between form and interpretation. Secondly, we fail to see any empirical or conceptual advantage in a system of syntactic heads solely using interpretively opaque elements such as Inflection rather than Tense or Aspect, Complementizer rather than Focus, Topic or Q marker, and so on. Conceptually, a transparent mapping surely is the null hypothesis, any deviation from which would require clear supporting evidence. Empirically, the transparent view is supported by much overt morphological evidence found across languages. Our own feeling is that the issue of cartography and the autonomy thesis should be looked at in the diametrically opposite perspective. The cartographic studies can be seen as an attempt to "syntacticize" as much as possible the interpretive domains, tracing back interpretive algorithms for such properties as argument structure (Hale and Keyser 1993 and much related work), scope, and informational structure (the "criterial" approach defended in Rizzi 1997 and much related work) to the familiar ingredients uncovered and refined in half a century of formal syntax. To the extent to which these efforts are empirically supported, they may shed light not only on syntax proper but also on the structure and functioning of the cognitive systems at the interface with the syntactic module.

3.6 Hierarchies, syntax, and semantics

Cartographic studies have drawn detailed structural maps holding across languages, and have made it plausible that core aspects of the functional structure may be universal. One important question which arises is: where does the hierarchy, and its universal properties, come from? It is hard to imagine that the hierarchy may be an irreducible property of UG, disconnected from any other aspect of human cognition; it is also hard to believe that the hierarchy may be a purely arbitrary "cultural" property, rediscovered by every language learner in the same form, language after language, on the basis of pure inductive learning. Therefore, there must be some principles determining the hierarchical sequence, and guiding the child to "rediscover" it in the course of language acquisition. In some cases, it is quite plausible that certain aspects of the hierarchy (such as the relative height, or scope, of the elements that constitute it) depend on independent properties of their semantics, even though precisely what elements make up the hierarchy may simply be the result of the linguistic crystallization of a particular set of cognitive categories among the many more that simply do not find a grammatical encoding in UG. Consider for instance the fact that many languages allow a proliferation of left-peripheral topics, while the left-peripheral focus position (if a language uses it at all) appears to be invariably unique. It is plausible that this difference may be derivable from the very interpretive properties of topic and focus (Rizzi 1997). If the left-peripheral focal head assigns the focus interpretation to its specifier and the presupposition interpretation to its complement,

(5) [XP [Foc YP]]
Focus Presupp.

then a recursion of (5), e.g., with YP headed by a Foc head, would yield an interpretive clash: YP would be presupposed, but would contain a focal constituent. Thus, the recursion is barred. On the other hand, nothing blocks the recursion of a topic phrase: no interpretive property of the comment rules out that it may in turn have a topic-comment structure. Individual languages may opt for a unique topic position as a matter of parametric choice, e.g., in V-2 languages, but there is no universal prohibition stemming from a plausible interpretive constraint in this case. Another example may be the fact that, in the structure of the IP, the element expressing epistemic modality typically is higher than tense: presumably the modality must be evaluated over a complete proposition, including the tense specification. Similar considerations may hold for the universal order epistemic modality > root modality, tense > aspects, etc.

In other cases, aspects of the hierarchy may be determined by syntactic constraints on movement. Consider for instance the fact that in many languages left-peripheral topic and focus can cooccur in the fixed order Topic–Focus (e.g., Hungarian: Kiss 1995). This may be due to the fact that Focus often requires movement of the inflected verb to C (possibly a property related to the quantificational character of Focus), while Topic does not. In a language requiring inversion with Focus, the order Focus–Topic would then be blocked by the impossibility of moving the inflected verb past the Topic head, ultimately a case of the Head Movement Constraint (Travis 1984). The validity of a syntactic account of this sort is supported by the fact that the order Focus–Topic seems indeed to be possible in a language like Italian, which does not require verb movement with focus. This strongly supports the view that in this case there is no general scope property enforcing a particular order. Along similar lines, one can observe that if a position has island-creating properties, it must be higher than other positions filled by movement: so, for instance, the Hanging Topic (which has island-creating properties) must precede the ordinary topic expressed in Romance by Clitic Left Dislocation (Cinque 1990; Benincà and Poletto 2004). On certain connections between the theory of movement and the hierarchy see Abels (2007).

Going back to the constraining effects of semantics, a qualification is needed. Clearly, it is not the case that any imaginable semantic property or distinction can be grammaticalized, expressed by a functional element, a special morphology, a morphosyntactic feature:[10] there is a fairly restrictive universal set of properties that can be expressed by the functional elements entering into the different hierarchies associated with clauses and phrases. Therefore, syntax is heavily constrained by semantics but is not totally malleable: on the one hand, semantics respects purely syntactic constraints (such as locality effects); on the other hand, it is often the case that a syntactic device has a core semantic function, but it often acquires an independent life of its own, as it were, extending its scope well beyond its core semantic function. Consider, for instance, grammatical gender, whose core function is the expression of natural gender, but which gets extended to express an arbitrary classification in the nominal lexicon; the expression of tense, situating the event in time with respect to the utterance time, but extending to become an obligatory property of the clausal hierarchy, so that also a tenseless mathematical or logical truth must be expressed via a tensed sentence; the subject–predicate articulation expressing the "aboutness" relation, but becoming a general, obligatory property of clausal structures, which forces the use of expletives if the event is not presented about a particular argument; etc. Syntax is organized to express meaning, but does not dissolve into the mere organization of meaningful units: UG expresses the possible items of the functional lexicon and the way in which they are organized into hierarchies, tailored on the needs of the expression of meanings, but not reducing to them.

[10] For example, in the extended projection of an NP, we find evidence for different types of quantifiers, demonstratives, numerals, for functional categories of diminutivization, numerical approximation, etc., but we never find expressed, it seems, distinctions relating to the magical or non-magical character of a number (as opposed to its approximation), nor specialized forms meaning dear-to-me, (dear-to-you), not-dear-to-me-and-you parallel to the universal demonstrative distinctions close-to-me, (close-to-you), not-close-to-me-and-you. One could easily multiply such theoretically possible, yet non-existing, functional distinctions (also see Cinque 1999: 224, fn.10 and related text).

This page intentionally left blank

CHAPTER 4

CATEGORIAL GRAMMAR

GLYN MORRILL

In categorial grammar expressions are classified by structured types. The set of types is the recursive closure of a set of atomic types under a number of type-constructors, so that types are algebraic terms like arithmetic expressions or logical formulas. The rules of categorial grammar are supposed to express the laws obeyed by the types according the "meaning" of the type-constructors. A set of such rules forms a deductive system or logical calculus. A categorial grammar consists of a system of types, a calculus for those types, and a lexicon which is an assignment of types to basic expressions. The language thereby defined consists of the expressions derivable from the lexicon by the calculus. In the purest form of categorial grammar, the calculus would express all and only the laws of the types, and would be universal. A grammar is then just a lexicon. Categorial grammar is thus highly lexicalist; in the ideal case, purely lexicalist. The syntactic configurations in which words appear are projected from their lexical types.

Ajdukiewicz (Ajdukiewicz, 1935) proposed to classify expressions by fractional types. Let there be basic types including N for name or (referring) nominal (e.g., *John, the man, the tall man that walks,...*), CN for count noun (e.g., *man, tall man that walks,...*), and S for statement or (declarative) sentence (e.g., *John sings, the man sings, the tall man that walks sings,...*). Then where A and B are types, $\frac{B}{A}$ is also a type, where we allow types within types recursively in both the denominator

Work partially funded by the DGICYT project TIN2005–08832–C03–03 (MOISES-BAR). Email morrill@lsi.upc.edu, http://www-lsi.upc.edu/~morrill/.

and the numerator. We call an expression of type $\frac{B}{A}$ a *functor* with A the *argument* type and B the *value* type; the idea is to say that the functor $\frac{B}{A}$ wants to combine with an A in order to form a B:

(1) $$\frac{B}{A}\ A \Rightarrow B$$

This resembles the arithmetic law of multiplication $\frac{B}{A} \times A = B$ and the logical law of modus ponens $A \to B, A \vdash B$. For *The man that walks sings* we can represent an analysis as follows:

(2) $$\dfrac{\dfrac{\dfrac{\textbf{the}}{\frac{N}{CN}} \quad \dfrac{\dfrac{\textbf{man}}{CN} \quad \dfrac{\dfrac{\textbf{that}}{\frac{(\frac{CN}{CN})}{(\frac{S}{N})}} \quad \dfrac{\textbf{walks}}{\frac{S}{N}}}{\frac{CN}{CN}}}{CN}}{N} \quad \dfrac{\textbf{sings}}{\frac{S}{N}}}{S}$$

The non-directional fractional divisions do not show where a functor seeks its argument. Bar-Hillel (Bar-Hillel, 1953) proposed to have directional divisions. Let us write $A\backslash B$ for a functor seeking its argument A to the immediate left and B/A for a functor seeking its argument A to the immediate right:

(3) $$A, A\backslash B \Rightarrow B$$
$$B/A, A \Rightarrow B$$

Then our example has the directional analysis:

(4) $$\dfrac{\dfrac{\dfrac{\textbf{the}}{N/CN} \quad \dfrac{\dfrac{\textbf{man}}{CN} \quad \dfrac{\dfrac{\textbf{that}}{(CN\backslash CN)/(N\backslash S)} \quad \dfrac{\textbf{walks}}{N\backslash S}}{CN\backslash CN}}{CN}}{N} \quad \dfrac{\textbf{sings}}{N\backslash S}}{S}$$

Ajdukiewicz allowed multiple-argument functors, e.g., $\frac{S}{NN}$. Let us then say in general that where A and B are types, $A \bullet B$ is a type, and that there is the type law:

(5) $$A, B \Rightarrow A \bullet B$$

4.1 Lambek calculus

Let us try to give a definite meaning to the type-constructors. Let a vocabulary V be a set of words. By V^+ we mean the set of all non-empty strings of these words. The types classify V^+, i.e., they stand for, or are interpreted as, subsets of V^+. Where P is an atomic type, let $[[P]]$ be its set of word-strings. Then $A \bullet B$ will be the set of all word-strings which are the concatenation of a string in A with a string in B. $A\backslash C$ will be the set of all word-strings which concatenate with any string in A on the left to form a C. C/B will be the set of all word-strings which concatenate with any string in B on the right to form a C. Formally:

(6) $$\begin{aligned} [[A \bullet B]] &= \{s_1 + s_2 \in V^+ | \ s_1 \in [[A]] \ \& \ s_2 \in [[B]]\} \\ [[A\backslash C]] &= \{s_2 \in V^+ | \text{ forall } s_1 \in [[A]], s_1 + s_2 \in [[C]]\} \\ [[C/B]] &= \{s_1 \in V^+ | \text{ forall } s_2 \in [[B]], s_1 + s_2 \in [[C]]\} \end{aligned}$$

Corresponding to this interpretation, the following laws of type-shift can now be verified as valid because in every interpretation the type on the left is a subset of the type on the right:

(7) $$\begin{aligned} &A \bullet (A\backslash C) \Rightarrow C \\ &(C/B) \bullet B \Rightarrow C \end{aligned}$$

But there are further valid type laws. If an expression is of type A it can combine with any expression of type $A\backslash C$ on its right to form a C, so for any expression, if it is of type A, it is also true to say that it is of type $C/(A\backslash C)$. Formally, $[[A]] \subseteq [[C/(A\backslash C)]]$ in every interpretation. Thus, the following law of type-shift, called lifting or type-raising or the Montague rule, is valid:

(8) $A \Rightarrow C/(A\backslash C)$

By symmetry, there is also a backward version:

(9) $B \Rightarrow (C/B)\backslash C$

In a similar way, for an expression to be of type C/B means that it concatenates with B's on the right to form C's. But that also means that it can concatenate with a B/D on the right, i.e., a B missing a D at its right edge, to form a C/D, i.e., a C also missing a D at its right edge. Formally, $[[C/B]] \subseteq [[(C/D)/(B/D)]]$ in every interpretation. Hence there is the following law of type-shift, called division or the Geach rule:

(10) $C/B \Rightarrow (C/D)/(B/D)$

The symmetric backward law is:

(11) $A\backslash C \Rightarrow (D\backslash A)\backslash(D\backslash C)$

$$A \Rightarrow A$$
reflexivity

$$A \bullet (B \bullet C) \Leftrightarrow (A \bullet B) \bullet C$$
associativity

$$B \Rightarrow A \backslash C \text{ iff } A \bullet B \Rightarrow C \text{ iff } A \Rightarrow C/B$$
residuation

$$A \Rightarrow C \text{ if } A \Rightarrow B \text{ and } B \Rightarrow C$$
transitivity

Figure 4.1. The categorical Lambek calculus

The question arises: what would be a calculus of all and only such valid laws? Consider the formal system shown in Figure 4.1, called the *categorical Lambek calculus* (Lambek, 1958). This calculus generates all and only the laws of type-shift which are valid according to the interpretation we have made explicit (Pentus, 1993).[1]

Given a Lambek categorial grammar lexicon, to determine whether a word-string is a sentence we need to establish whether it can be factored into basic expressions $w_0, \ldots, w_n$ of lexical types $A_0, \ldots, A_n$ respectively such that $A_0 \bullet \cdots \bullet A_n \Rightarrow S$ is a valid type law (with any parenthesization on the left, because the product is associative), so we want a procedure to decide this question. The categorical calculus is not a good basis for trying to do this because to test whether $A \Rightarrow C$ can be derived by transitivity we need to determine whether there exists a type B such that $A \Rightarrow B$ and $B \Rightarrow C$ where B can be any of the infinite number of types, and we cannot try them all. A decision procedure can be provided however on the basis of another calculus called the *Lambek sequent calculus*.

Let a *sequent* $A_0, \ldots, A_n \Rightarrow A$ comprise a *succedent* type A and an *antecedent* configuration $A_0, \ldots, A_n$ which is a finite non-empty sequence of types. The *Lambek sequent calculus* is as shown in Figure 4.2, where $\Delta(\Gamma)$ indicates a configuration Δ with a distinguished subconfiguration Γ. Observe that for each connective there is a left (L) rule introducing it in the antecedent, and a right (R) rule introducing it in the succedent.

A sequent $A_0, \ldots, A_n \Rightarrow A$ is derivable in the sequent calculus if and only if $A_0 \bullet \cdots \bullet A_n \Rightarrow A$ (with any parenthesization of the products) is derivable in the categorical calculus (Lambek, 1958). Computationally, the Cut rule of the sequent calculus is problematic in the same way as the transitivity rule of the categorical calculus, introducing a new unknown (A) reading from conclusion to premise.

[1] In logical jargon, it is sound (every derivable arrow $A \Rightarrow B$ is valid) and complete (every valid arrow $A \Rightarrow B$ is derivable).

$$\frac{}{A \Rightarrow A}\,id \qquad \frac{\Gamma \Rightarrow A \qquad \Delta(A) \Rightarrow B}{\Delta(\Gamma) \Rightarrow B}\,Cut$$

$$\frac{\Gamma \Rightarrow A \qquad \Delta(C) \Rightarrow D}{\Delta(\Gamma, A\backslash C) \Rightarrow D}\,\backslash L \qquad \frac{A, \Gamma \Rightarrow C}{\Gamma \Rightarrow A\backslash C}\,\backslash R$$

$$\frac{\Gamma \Rightarrow B \qquad \Delta(C) \Rightarrow D}{\Delta(C/B, \Gamma) \Rightarrow D}\,/L \qquad \frac{\Gamma, B \Rightarrow C}{\Gamma \Rightarrow C/B}\,/R$$

$$\frac{\Delta(A, B) \Rightarrow D}{\Delta(A \bullet B) \Rightarrow D}\,\bullet L \qquad \frac{\Gamma \Rightarrow A \qquad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A \bullet B}\,\bullet R$$

Figure 4.2. Lambek sequent calculus

However, Lambek (1958) proved *Cut-elimination* for the sequent calculus: that every derivable sequent has a Cut-free derivation. All the other rules have the property that the number of type-constructor occurrences in the premises is one less than that in the conclusion. Cut-free backward chaining (reading from conclusions to premises) sequent proof search therefore operates in a finite space and constitutes a decision procedure for Lambek theoremhood.

In Cut-free sequent calculus, the example (4) can be derived:

(12)
$$\dfrac{\dfrac{\dfrac{N \Rightarrow N \qquad S \Rightarrow S}{N, N\backslash S \Rightarrow S}\,\backslash L}{N\backslash S \Rightarrow N\backslash S}\,\backslash R \qquad \dfrac{\dfrac{CN \Rightarrow CN \qquad \dfrac{CN \Rightarrow CN \qquad N \Rightarrow N}{N/CN, CN \Rightarrow N}\,/L}{N/CN, CN, CN\backslash CN \Rightarrow N}\,\backslash L \qquad S \Rightarrow S}{N/CN, CN, CN\backslash CN, N\backslash S \Rightarrow S}\,\backslash L}{N/CN, CN, (CN\backslash CN)/(N\backslash S), N\backslash S, N\backslash S \Rightarrow S}\,/L$$

Assuming a finite lexicon, to determine whether a word-string is a sentence we may apply the backward chaining Cut-free sequent decision procedure to each of the finite number of sequents $A_0, \ldots, A_n \Rightarrow S$ into which it can be lexically analyzed.

4.2 Basic Discontinuous Lambek Calculus

The Lambek calculus captures the laws of the categorial type-constructors $\backslash, \bullet, /$ interpreted with respect to concatenation. By *discontinuous Lambek calculus* we mean calculus that aspires to extend to *interpolation* what the Lambek calculus is for concatenation, i.e., to extend to discontinuity what the Lambek calculus does for continuity.

Let us assume that strings over the vocabulary may include also a special new place-holder symbol 1 called the *separator*, and that we also now allow the empty string. We define the sort of a string as the number of separators it contains: for each natural $n \geq 0$ we define the set L_n of strings of sort n as the set of all strings over the vocabulary and separator which contain n separators.

Here we consider a minimal discontinuous Lambek calculus, which we call Basic Discontinuous Lambek Calculus (BDLC) (Morrill and Fadda, 2008), and which involves only the two sorts 0 and 1. In addition to defining continuous operators $\backslash, \bullet, /$ with respect to concatenation + of functionality $L_0 \times L_0 \rightarrow L_0$, we define discontinuous operators $\downarrow, \odot, \uparrow$ with respect to interpolation or wrapping W of functionality $L_1 \times L_0 \rightarrow L_0$ such that $(s_1 + 1 + s_3)Ws_2 = s_1 + s_2 + s_3$:

(13)
$$\begin{aligned}
[[A \bullet B]] &= \{s_1 + s_2 \in L_0 |\ s_1 \in [[A]]\ \&\ s_2 \in [[B]]\} \\
[[A \backslash C]] &= \{s_2 \in L_0 |\ \text{forall}\ s_1 \in [[A]], s_1 + s_2 \in [[C]]\} \\
[[C / B]] &= \{s_1 \in L_0 |\ \text{forall}\ s_2 \in [[B]], s_1 + s_2 \in [[C]]\} \\
[[A \odot B]] &= \{s_1 W s_2 \in L_0 |\ s_1 \in [[A]]\ \&\ s_2 \in [[B]]\} \\
&= \{s_1 + s_2 + s_3 \in L_0 |\ s_1 + 1 + s_3 \in [[A]]\ \&\ s_2 \in [[B]]\} \\
[[A \downarrow C]] &= \{s_2 \in L_0 |\ \text{forall}\ s_1 \in [[A]], s_1 W s_2 \in [[C]]\} \\
&= \{s_2 \in L_0 |\ \text{forall}\ s_1 + 1 + s_3 \in [[A]], s_1 + s_2 + s_3 \in [[C]]\} \\
[[C \uparrow B]] &= \{s_1 \in L_1 |\ \text{forall}\ s_2 \in [[B]], s_1 W s_2 \in [[C]]\} \\
&= \{s_1 + 1 + s_3 \in L_1 |\ \text{forall}\ s_2 \in [[B]], s_1 + s_2 + s_3 \in [[C]]\}
\end{aligned}$$

We define types $\mathcal{F}_0$ of sort 0 and $\mathcal{F}_1$ of sort 1 as follows:

(14)
$$\begin{aligned}
\mathcal{F}_0 &::= \mathcal{A}_0 \mid \mathcal{F}_0 \backslash \mathcal{F}_0 \mid \mathcal{F}_0 / \mathcal{F}_0 \mid \mathcal{F}_0 \bullet \mathcal{F}_0 \mid \mathcal{F}_1 \downarrow \mathcal{F}_0 \mid \mathcal{F}_1 \odot \mathcal{F}_0 \\
\mathcal{F}_1 &::= \mathcal{F}_0 \uparrow \mathcal{F}_0
\end{aligned}$$

An expression of a discontinuous (sort 1) type is of the form $s_1 + 1 + s_2$. Because of the presence of interpolation as well as concatenation, the two components s_1 and s_2 may become separated by an infix and the separator may disappear. However, s_1 will always remain to the left of s_2, and because interpolation retains the infix intact, the components of discontinuous types will always stay well-nested, i.e., well-bracketed as open- and close-parentheses.

To give sequent calculus for discontinuous Lambek calculus we represent a sort 1 discontinuous type A in a sequent by two tokens $\sqrt[0]{A}$ and $\sqrt[1]{A}$ at the positions of its first and second components respectively. We represent the separator by [ ]. A continuous (sort 0) type A is represented in antecedents and succedents as a single token A. A discontinuous (sort 1) type A is represented in a succedent as $\sqrt[0]{A}$, [ ], $\sqrt[1]{A}$. A discontinuous type A in an antecedent is represented by $\sqrt[0]{A}$ to the left of $\sqrt[1]{A}$; as has been remarked, all the respective roots are well-nested; if the succedent is of sort 1, the metalinguistic separator [ ] appears at some point in

$$\frac{}{A \Rightarrow A}\ id$$

$$\frac{\Gamma \Rightarrow A \qquad \Delta(A) \Rightarrow X}{\Delta(\Gamma) \Rightarrow X}\ Cut_0$$

$$\frac{\Gamma([\,]) \Rightarrow \sqrt[0]{A}, [\,], \sqrt[1]{A} \qquad \Delta(\sqrt[0]{A}, \Omega, \sqrt[1]{A}) \Rightarrow X}{\Delta(\Gamma(\Omega)) \Rightarrow X}\ Cut_1$$

$$\frac{\Gamma \Rightarrow A \qquad \Delta(C) \Rightarrow X}{\Delta(\Gamma, A\backslash C) \Rightarrow X}\ \backslash L \qquad \frac{A, \Gamma \Rightarrow C}{\Gamma \Rightarrow A\backslash C}\ \backslash R$$

$$\frac{\Gamma \Rightarrow B \qquad \Delta(C) \Rightarrow X}{\Delta(C/B, \Gamma) \Rightarrow X}\ /L \qquad \frac{\Gamma, B \Rightarrow C}{\Gamma \Rightarrow C/B}\ /R$$

$$\frac{\Delta(A, B) \Rightarrow X}{\Delta(A\bullet B) \Rightarrow X}\ \bullet L \qquad \frac{\Gamma \Rightarrow A \qquad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A\bullet B}\ \bullet R$$

$$\frac{\Gamma([\,]) \Rightarrow \sqrt[0]{A}, [\,], \sqrt[1]{A} \qquad \Delta(C) \Rightarrow X}{\Delta(\Gamma(A{\downarrow}C)) \Rightarrow X}\ {\downarrow}L \qquad \frac{\sqrt[0]{A}, \Gamma, \sqrt[1]{A} \Rightarrow C}{\Gamma \Rightarrow A{\downarrow}C}\ {\downarrow}R$$

$$\frac{\Gamma \Rightarrow B \qquad \Delta(C) \Rightarrow X}{\Delta(\sqrt[0]{C{\uparrow}B}, \Gamma, \sqrt[1]{C{\uparrow}B}) \Rightarrow X}\ {\uparrow}L \qquad \frac{\Gamma(B) \Rightarrow C}{\Gamma([\,]) \Rightarrow C{\uparrow}B}\ {\uparrow}R$$

$$\frac{\Delta(\sqrt[0]{A}, B, \sqrt[1]{A}) \Rightarrow X}{\Delta(A\odot B) \Rightarrow X}\ \odot L \qquad \frac{\Gamma([\,]) \Rightarrow \sqrt[0]{A}, [\,], \sqrt[1]{A} \qquad \Delta \Rightarrow B}{\Gamma(\Delta) \Rightarrow A\odot B}\ \odot R$$

Figure 4.3. BDLC hypersequent calculus

the antecedent. Thus the sets $\mathcal{O}_0$ and $\mathcal{O}_1$ of *configurations of sort 0 and 1* are defined as follows:

(15) $\mathcal{O}_0 ::= \Lambda \mid A_0, \mathcal{O}_0 \mid \sqrt[0]{A_1}, \mathcal{O}_0, \sqrt[1]{A_1}, \mathcal{O}_0$
$\mathcal{O}_1 ::= \mathcal{O}_0, [\,], \mathcal{O}_0 \mid \mathcal{O}_0, \sqrt[0]{A_1}, \mathcal{O}_1, \sqrt[1]{A_1}, \mathcal{O}_0$

The sequent calculus of this kind, which we call *hypersequent calculus*, for the Basic Discontinuous Lambek Calculus (BDLC) is shown in Figure 4.3, where X ranges over sort 0 and sort 1 succedents.

By way of example of discontinuity, consider that under "pied-piping" a relative pronoun may occur embedded:

(16) a. the painting which won
b. the bicycle the owner of which whistles

The lexical assignment of a single relative pronoun type $(N \uparrow N) \downarrow ((CN\backslash CN)/(N\backslash S))$ optionally allows the embedding. Thus (16a) is derived thus:

$$(17)\quad \dfrac{\dfrac{N \Rightarrow N}{[\,] \Rightarrow \sqrt[0]{N{\uparrow}N}, [\,], \sqrt[1]{N{\uparrow}N}}\,{\uparrow}R \qquad \dfrac{\dfrac{\dfrac{N \Rightarrow N \qquad S \Rightarrow S}{N, N\backslash S \Rightarrow S}\,\backslash L}{N\backslash S \Rightarrow N\backslash S}\,\backslash R \qquad \dfrac{CN \Rightarrow CN \qquad \dfrac{CN \Rightarrow CN \qquad N \Rightarrow N}{N/CN, CN \Rightarrow N}\,/L}{N/CN, CN, CN\backslash CN \Rightarrow N}\,\backslash L}{N/CN, CN, (CN\backslash CN)/(N\backslash S), N\backslash S \Rightarrow N}\,/L}{N/CN, CN, (N{\uparrow}N){\downarrow}((CN\backslash CN)/(N\backslash S)), N\backslash S \Rightarrow N}\,{\downarrow}L$$

And (16b) is derived as follows (which may be completed by the reader):

$$(18)\quad \dfrac{\dfrac{N/CN, CN/PP, PP/N, N \Rightarrow N}{N/CN, CN/PP, PP/N, [\,] \Rightarrow \sqrt[0]{N{\uparrow}N}, [\,], \sqrt[1]{N{\uparrow}N}}\,{\uparrow}R \qquad \dfrac{N\backslash S \Rightarrow N\backslash S \qquad N/CN, CN, CN\backslash CN \Rightarrow N}{N/CN, CN, (CN\backslash CN)/(N\backslash S), N\backslash S \Rightarrow N}\,/L}{N/CN, CN, N/CN, CN/PP, PP/N, (N{\uparrow}N){\downarrow}((CN\backslash CN)/(N\backslash S)), N\backslash S \Rightarrow N}\,{\downarrow}L$$

4.3 Syntactic structures as proof nets

Logical calculi in general, and categorial calculi in particular, can be presented in many formats: sequent calculus, natural deduction, categorical calculus, and so on. In categorial grammar, as we construe it here, derivations are proofs. So syntactic structures are to be equated with the structure of proofs. But according to which proof format? We want syntactic structures to be canonical, i.e., to be unique for each semantic reading. And we want them to be as free as possible from redundancy and unnecessary clutter. The sequent calculus, for example, is both non-canonical in that there may be multiple sequent derivations delivering the same reading, and redundant in that there is much copying from premises to conclusions at every step. Roorda (Roorda, 1991) observes that the linear logic proof nets of Girard (1987) can be adapted to categorial grammar to neatly fulfill the role of canonical and non-redundant syntactic structure.

We formulate proof nets for the Lambek calculus and for BDLC in a variant of the "highway" notation of Morrill and Fadda (2008), in which the branches of type trees are single or multiple lanes.

A *polar type* A^p comprises a type A and a polarity $p = \bullet$ (input) or $\circ$ (output). We define the *complements* $\overline{A^\bullet} =_{df.} A^o$ and $\overline{A^o} =_{df.} A^\bullet$. The continuous (Lambek) *logical links* are as shown in Figure 4.4. The intermediate nodes $\otimes$ and $\wp$ represent the (multiplicative) linear logic conjunction and disjunction respectively underlying each link. We refer to lane edges as *parameter edges* and we refer to an uninterrupted sequence of dashed parameter edges as a $\forall$-*segment*. We refer to entire, possibly multilane, highways seen as single broad edges as *predicate edges*.

A *polar type tree* is the result of unfolding a polar type up to its atomic leaves according to the logical links. A *proof frame* for a sequent $A_0, \ldots, A_n \Rightarrow A$ is the

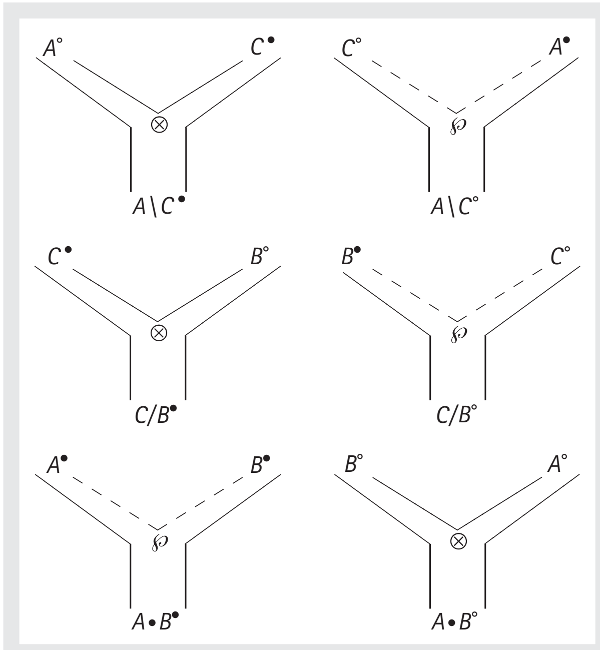
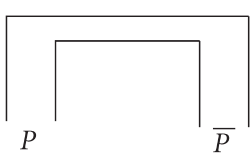


Figure 4.4. Continuous logical links

multiset of unfolded polar type trees of A^o, $A_1{}^\bullet, \ldots, A_n{}^\bullet$. An *axiom link* is as follows, where P is an atomic polar type:

(19)



A *proof structure* is a proof frame to which have been added axiom links connecting each leaf to exactly one other complementary leaf. A *proof net* is a proof structure satisfying the following correctness criteria:

(20)
- (*Danos-Regnier acyclicity*) Every predicate edge cycle crosses both premise edges of some $\wp$-link.
- ($\forall$-*correctness*) Every parameter edge cycle contains exactly one $\forall$-segment and if a parameter path does not form part of a cycle then it does not contain any $\forall$-segment.

The highway proof net syntactic structure for our example *The man that walks sings* is shown in Figure 4.5. It contains one predicate edge cycle, which respects Danos-Regnier acyclicity, and the associated parameter edge cycle respects $\forall$-correctness. The proof net is planar (it can be drawn in the half plane without crossing edges) as are all Lambek proof nets, but not all BDLC proof nets.

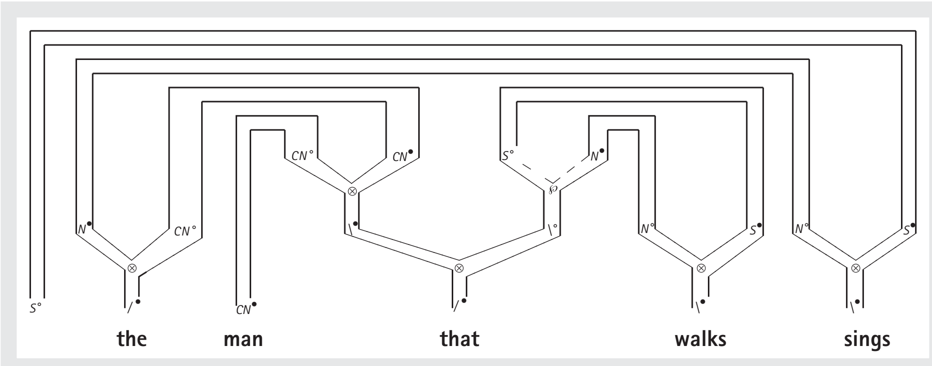


Figure 4.5. Proof net syntactic structure for *The man that walks sings*

The logical links for the discontinuous connectives of BDLC are given in Figure 4.6. For Basic Discontinuous Lambek Calculus we need to augment the correctness criteria (20), but apparently the criterion of Morrill and Fadda (2008) is insufficient.

Henceforth we shall represent proof net syntactic structures in outline only, presenting just polar type trees with predicate edges as single edges and without

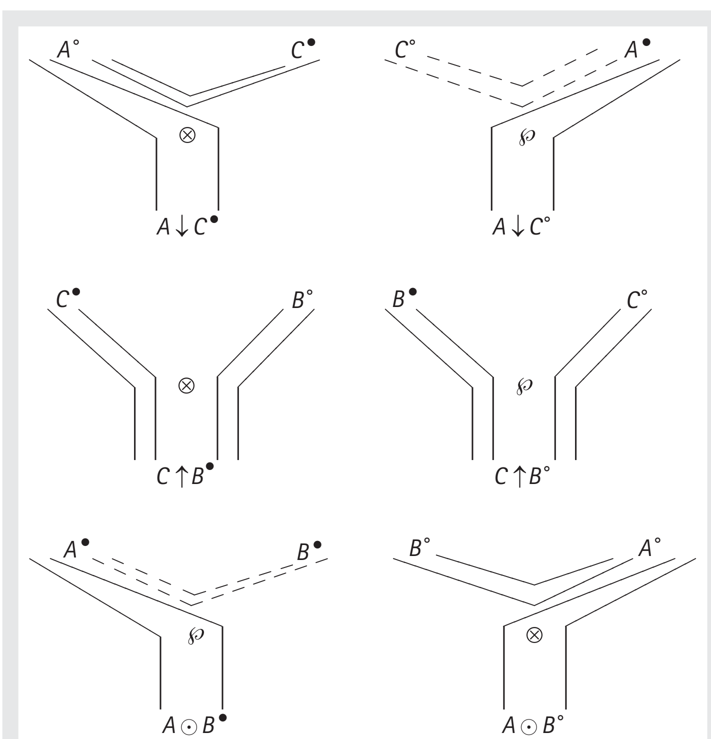


Figure 4.6. Discontinuous logical links

intermediate nodes. These contain all the information necessary to recover the full expanded highway structures to which the correctness criteria refer, but which would clutter and obscure what is essential to our present purposes. By way of example, Figure 4.7 contains the outline proof net syntactic structure of (16a), corresponding to the sequent derivation (17). And Figure 4.8 presents, with some abbreviation, the outline proof net syntactic structure of (16b), corresponding to the sequent derivation started in (18).

4.4 Semantics

Categorial semantics derives from the reading of categorial derivations as intuitionistic proofs.

4.4.1 Intuitionistic sequent calculus

Here, let us define a *sequent* $\Gamma \Rightarrow A$ as comprising an *antecedent* Γ which is a finite sequence of formulas, and a *succedent* formula A. We read a sequent as asserting that the conjunction of the antecedent formulas entails the succedent formula. A sequent is called *valid* if and only if this assertion is true; otherwise it is called *invalid*. A sequent calculus for implicative and conjunctive intuitionistic logic is presented in Figure 4.9.

The rules W (weakening), C (contraction) and P (permutation) are referred to as *structural* rules; they apply to properties of all formulas with respect to the metalinguistic comma interpreted as conjunction in the antecedent. Note that these rules are absent in the Lambek calculus (and in discontinuous Lambek calculus).

4.4.2 Natural deduction

Natural deduction is a single-conclusioned proof format particularly suited to intuitionistic logic. A natural deduction proof is a tree of formulas with some coindexing of leaves with dominating nodes. The leaf formulas are called *hypotheses: open* if not indexed, *closed* if indexed. The root of the tree is the *conclusion*: a natural deduction proof asserts that (any superset of) its open hypotheses entail(s) its conclusion. A trivial tree consisting of a single formula is a proof (from itself, as open hypothesis, to itself, as conclusion, corresponding to the identity axiom *id* of sequent calculus). Then the proofs of $\{\rightarrow, \wedge\}$-intuitionistic logic are those further generated by the rules in Figure 4.10. Note that hypotheses become indexed (closed) when the dominating coindexed inference occurs, and any number of hypotheses (including zero) can be indexed/closed in one step, cf. weakening and contraction.

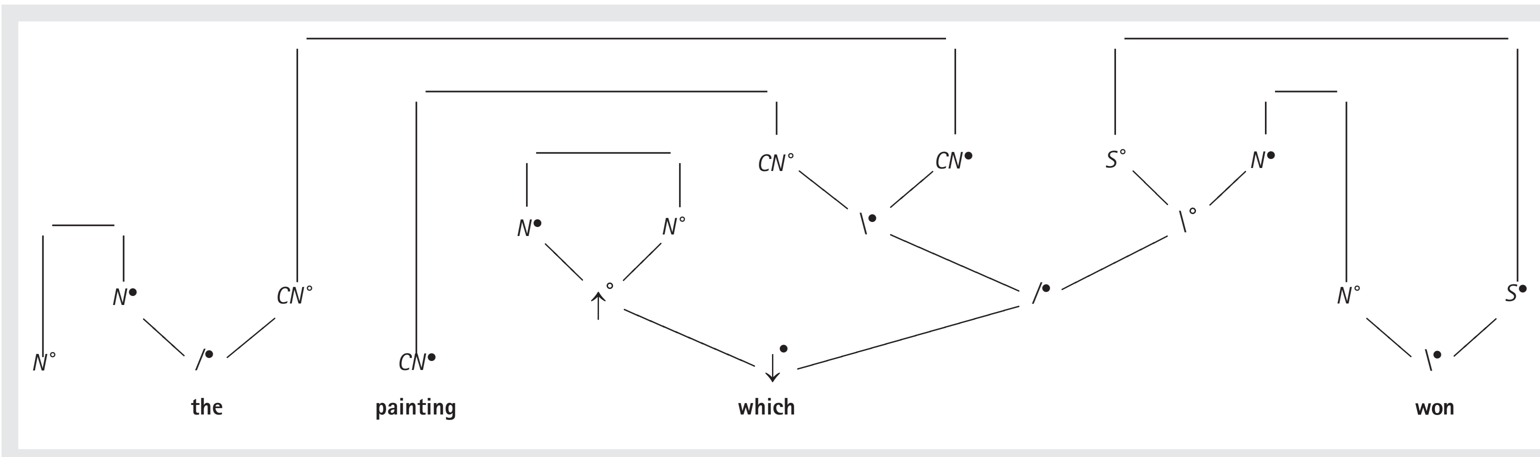


Figure 4.7. Outline proof net syntactic structure for *the painting which won*

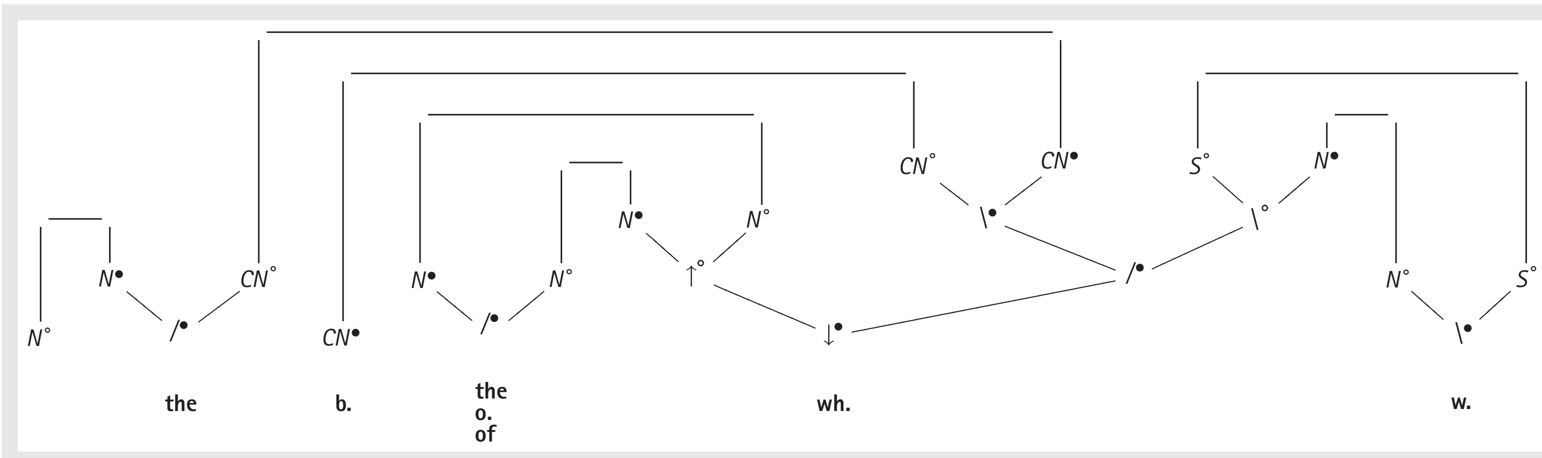


Figure 4.8. Outline proof net syntactic structure for *the bicycle the owner of which whistles*

$$\frac{}{A \Rightarrow A}\ id \qquad \frac{\Gamma \Rightarrow A \qquad \Delta_1, A, \Delta_2 \Rightarrow B}{\Delta_1, \Gamma, \Delta_2 \Rightarrow B}\ Cut$$

$$\frac{\Delta_1, \Delta_2 \Rightarrow C}{\Delta_1, A, \Delta_2 \Rightarrow C}\ W$$

$$\frac{\Delta_1, A, A, \Delta_2 \Rightarrow C}{\Delta_1, A, \Delta_2 \Rightarrow C}\ C$$

$$\frac{\Delta_1, A, B, \Delta_2 \Rightarrow C}{\Delta_1, B, A, \Delta_2 \Rightarrow C}\ P$$

$$\frac{\Delta_1, A, B, \Delta_2 \Rightarrow C}{\Delta_1, A \wedge B, \Delta_2 \Rightarrow C}\ \wedge L \qquad \frac{\Delta_1 \Rightarrow A \qquad \Delta_2 \Rightarrow B}{\Delta_1, \Delta_2 \Rightarrow A \wedge B}\ \wedge R$$

$$\frac{\Gamma \Rightarrow A \qquad \Delta_1, B, \Delta_2 \Rightarrow C}{\Delta_1, \Gamma, A \rightarrow B, \Delta_2 \Rightarrow C}\ \rightarrow L \qquad \frac{\Delta_1, A, \Delta_2 \Rightarrow B}{\Delta_1, \Delta_2 \Rightarrow A \rightarrow B}\ \rightarrow R$$

Figure 4.9. Sequent calculus for $\{\rightarrow, \wedge\}$-intuitionistic logic

4.4.3 Typed lambda calculus

The untyped lambda calculus was introduced as a model of computation by Alonzo Church. It uses a variable binding operator (the λ) to name functions, and forms the basis of functional programming languages such as LISP. It was proved equivalent to Turing machines, hence the name Church-Turing Thesis for the notion that Turing machines (and untyped lambda calculus) capture the notion of algorithm. Church

$$\frac{\begin{array}{cc}\vdots & \vdots \\ A & A \rightarrow B\end{array}}{B}\ E \rightarrow \qquad \frac{\begin{array}{c}A^{i} \\ \vdots \\ B\end{array}}{A \rightarrow B}\ I \rightarrow^{i}$$

$$\frac{\begin{array}{c}\vdots \\ A \wedge B\end{array}}{A}\ E\wedge_1 \qquad \frac{\begin{array}{c}\vdots \\ A \wedge B\end{array}}{B}\ E\wedge_2 \qquad \frac{\begin{array}{cc}\vdots & \vdots \\ A & B\end{array}}{A \wedge B}\ I\wedge$$

Figure 4.10. Natural deduction rules for $\{\rightarrow, \wedge\}$-intuitionistic logic

(1940) defined the simply, i.e., just functionally, typed lambda calculus, and, by including logical constants, higher-order logic. Here we add also Cartesian product types.

(21) **Definition** (*types*)
The T set of *types* is defined on the basis of a set δ of *basic types* as follows:

$$\mathcal{T} ::= \delta | \mathcal{T} \rightarrow \mathcal{T} | \mathcal{T} \& \mathcal{T}$$

(22) **Definition** (*type domains*)
The *type domain* D_τ of each type τ is defined on the basis of an assignment d of non-empty sets (*basic type domains*) to δ as follows:

$D_\tau = d(\tau)$	for $\tau \in \delta$
$D_{\tau_1 \rightarrow \tau_2} = D_{\tau_2}^{D_{\tau_1}}$	functional exponentiation
	i.e., the set of all functions from D_{τ_1} to D_{τ_2}
$D_{\tau_1 \& \tau_2} = D_{\tau_1} \times D_{\tau_2}$	Cartesian product
	i.e., $\{\langle m_1, m_2 \rangle | m_1 \in D_{\tau_1} \& m_2 \in D_{\tau_2}\}$

(23) **Definition** (*terms*)
The sets Φ_τ of *terms* of type τ for each type τ are defined on the basis of a set C_τ of constants of type τ and a denumerably infinite set V_τ of variables of type τ for each type τ as follows:

$$\begin{aligned} \Phi_\tau &::= C_\tau \mid V_\tau \mid (\Phi_{\tau' \rightarrow \tau}\ \Phi_{\tau'}) \mid \pi_1 \Phi_{\tau \& \tau'} \mid \pi_2 \Phi_{\tau' \& \tau} \\ \Phi_{\tau \rightarrow \tau'} &::= \lambda V_\tau \Phi_{\tau'} \\ \Phi_{\tau \& \tau'} &::= (\Phi_\tau, \Phi_{\tau'}) \end{aligned}$$

Each term $\phi \in \Phi_\tau$ receives a semantic value $[\phi]^g \in D_\tau$ with respect to a valuation f which is a mapping sending each constant in C_τ to an element in D_τ, and an assignment g which is a mapping sending each variable in V_τ to an element in D_τ, as shown in Figure 4.11.

$[c]^g = f(c)$	for $c \in C_\tau$
$[x]^g = g(x)$	for $x \in V_\tau$
$[(\phi\ \psi)]^g = [\phi]^g([\psi]^g)$	functional application
$[\pi_1\ \phi]^g = \mathsf{fst}([\phi]^g)$	first projection
$[\pi_2\ \phi]^g = \mathsf{snd}([\phi]^g)$	second projection
$[\lambda x_\tau \phi]^g = D_\tau \ni d \mapsto [\phi]^{(g - \{(x, g(x))\}) \cup \{(x, d)\}}$	functional abstraction
$[(\phi, \psi)]^g = \langle [\phi]^g, [\psi]^g \rangle$	ordered pair formation

Figure 4.11. Semantics of typed lambda calculus

$$\lambda y\phi = \lambda x(\phi\{x/y\})$$
if x is not free in ϕ and x is free for y in ϕ

α-conversion

$$(\lambda x\phi\ \psi) = \phi\{\psi/x\}$$
if ψ is free for x in ϕ
$$\pi_1(\phi, \psi) = \phi$$
$$\pi_2(\phi, \psi) = \psi$$

β-conversion

$$\lambda x(\phi\ x) = \phi$$
if x is not free in ϕ
$$(\pi_1\phi, \pi_2\phi) = \phi$$

η-conversion

Figure 4.12. Laws of lambda conversion

An occurrence of a variable in a term is called *free* if and only if it does not fall within any part of the term of the form $\lambda x\cdot$; otherwise it is *bound* (by the closest λx within the scope of which it falls). The result $\phi\{\psi/x\}$ of substituting term ψ (of type τ) for variable x (of type τ) in a term ϕ is the result of replacing by ψ every free occurrence of x in ϕ. We say that ψ is *free for x in ϕ* if and only if no variable in ψ becomes bound in $\phi\{\psi/x\}$. Manipulations may be pathological if substitution is not free in this sense. The laws of lambda conversion in Figure 4.12 obtain.

The Curry-Howard correspondence (Girard et al., 1989) is that intuitionistic natural deduction and typed lambda calculus are isomorphic. This *formulas-as-types* and *proofs-as-programs* correspondence exists at the following three levels:

(24)

intuitionistic natural deduction	typed lambda calculus
formulas:	types:
$A \rightarrow B$	$\tau_1 \rightarrow \tau_2$
$A \wedge B$	$\tau_1 \& \tau_2$
proofs:	terms:
E(limination of) $\rightarrow$	functional application
I(ntroduction of) $\rightarrow$	functional abstraction
E(limination of) $\wedge$	projection
I(ntroduction of) $\wedge$	ordered pair formation
normalization:	computation:
elimination of detours	lambda reduction

$$\dfrac{\dfrac{\begin{matrix}\phi & \psi\\ \vdots & \vdots\\ A & B\end{matrix}}{A \wedge B}\, I\wedge_1}{A}\, E\wedge \quad \rightsquigarrow \quad \begin{matrix}\phi\\ \vdots\\ A\end{matrix} \qquad\qquad \dfrac{\dfrac{\begin{matrix}\phi & \psi\\ \vdots & \vdots\\ A & B\end{matrix}}{A \wedge B}\, I\wedge_2}{B}\, E\wedge \quad \rightsquigarrow \quad \begin{matrix}\psi\\ \vdots\\ B\end{matrix}$$

Figure 4.13. β-reduction for conjunction

$$\dfrac{\dfrac{\begin{matrix}\phi\\ \vdots\\ A \wedge B\end{matrix}}{A}\, E\wedge_1 \qquad \dfrac{\begin{matrix}\phi\\ \vdots\\ A \wedge B \quad 1\end{matrix}}{B}\, E\wedge_2}{A \wedge B}\, I\wedge \quad \rightsquigarrow \quad \begin{matrix}\phi\\ \vdots\\ A \wedge B\end{matrix}$$

Figure 4.14. η-reduction for conjunction

Overall, the laws of lambda reduction are the same as the natural deduction proof normalizations (Prawitz, 1965). By way of illustration, the β- and η-proof reductions for conjunction are as shown in Figures 4.13 and 4.14 respectively. In contrast to the untyped lambda calculus, the normalization of terms (evaluation of "programs") in the typed lambda calculus is *terminating*: every term reduces to a normal form in a finite number of steps.

4.4.4 Type-logical semantics

Our categorial proofs are valid as intuitionistic proofs under the following type map:

(25)
$$\begin{aligned} T(A\backslash C) &= T(A) \to T(B)\\ T(C/B) &= T(B) \to T(C)\\ T(A \bullet B) &= T(A) \wedge T(B)\\ T(A \downarrow C) &= T(A) \to T(B)\\ T(C \uparrow B) &= T(B) \to T(C)\\ T(A \odot B) &= T(A) \wedge T(B) \end{aligned}$$

It follows, under the Curry-Howard correspondence, that every categorial derivation is associated with a typed lambda term, which we take as the derivational semantics. Lexical semantics (closed terms) is substituted into the derivational semantics (an open term) to obtain the semantics of the derived expression (a closed term). We call this architecture "Curry-Howard type-logical semantics".

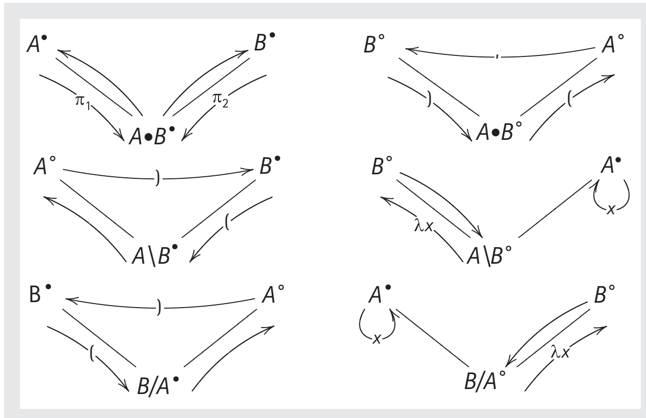


Figure 4.15. Continuous semantic trip travel instructions

When derivations are represented as proof nets, the semantic reading is extracted by a special trip around the proof net which we call the "semantic trip". The semantic trip begins at the unique root of polarity output (the "origin"), and starts by traveling upward. Edges are followed in uniform direction until we come to logical links. Then travel instructions are followed as given in Figure 4.15 for the continuous connectives and in Figure 4.16 for the discontinuous connectives. The labels of edges taken generate the successive symbols of the semantic form. Lambda variables are unique to their link. When we arrive down at an input polarity root, the associated lexical semantics is inserted, and the trip "bounces" back up. The trip visits each node twice, once traveling upward and once traveling downward, and crosses each edge twice, once in each direction. It ends when it arrives back down at the origin.

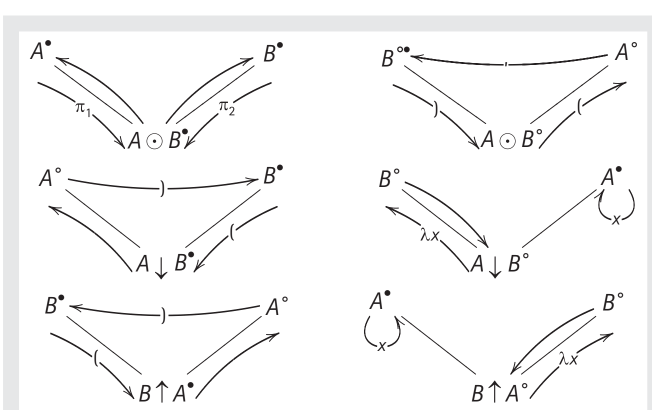


Figure 4.16. Discontinuous semantic trip travel instructions

For our initial example, for which a proof net analysis is given in Figure 4.5, let there be lexical semantics as follows:

(26)
$$\begin{aligned}
\textbf{man} &- \text{man}\\
&:= CN\\
\textbf{sings} &- \text{sing}\\
&:= N\backslash S\\
\textbf{that} &- \lambda x\lambda y\lambda z((\wedge\ (y\ z))\ (x\ z))\\
&:= (CN\backslash CN)/(N\backslash S)\\
\textbf{the} &- \iota\\
&:= N/CN\\
\textbf{walks} &- \text{walk}\\
&:= N\backslash S
\end{aligned}$$

The result of the semantic trip is (27a), which normalizes to (27b).

(27) a. $(sing\ (\iota\ ((\lambda x\lambda y\lambda z((\wedge\ (y\ z))\ (x\ z))\ \lambda w(walk\ w))\ man)))$
b. $(sing\ (\iota\ \lambda z((\wedge\ (man\ z))\ (walk\ z))))$

4.5 Complexity

Finally, let us consider quantifier scoping and quantifier scope preference. Let there be lexical assignments as follows:

(28)
$$\begin{aligned}
\textbf{everyone} &- \lambda x(\forall\ \lambda y((\rightarrow\ (person\ y))\ (x\ y)))\\
&:= (S \uparrow N) \downarrow S\\
\textbf{loves} &- love\\
&:= (N\backslash S)/N\\
\textbf{someone} &- \lambda x(\exists\ \lambda y((\wedge\ (person\ y))\ (x\ y)))\\
&:= (S \uparrow N) \downarrow S
\end{aligned}$$

Such assignments allow quantifier phrases to occupy nominal positions while taking semantic scope sententially. For *Everyone loves someone* there are the proof net analyses of Figure 4.17 and Figure 4.18 with the subject-wide scope semantics (29a) and the object-wide scope semantics (29b) respectively.

(29) a. $(\forall\ \lambda x((\rightarrow\ (person\ x))\ (\exists\ \lambda y((\wedge\ (person\ y))\ ((love\ y)\ x)))))$
b. $(\exists\ \lambda y((\wedge\ (person\ y))\ (\forall\ \lambda x((\rightarrow\ (person\ x))\ ((love\ y)\ x)))))$

Morrill (2000) proposes for a range of performance phenomena a simple metric of complexity profile of proof net analyses, this being the graph showing the number of unresolved dependencies at each word boundary assuming an incremental (i.e., left-to-right) construction of proof net syntactic structures. We interpret this

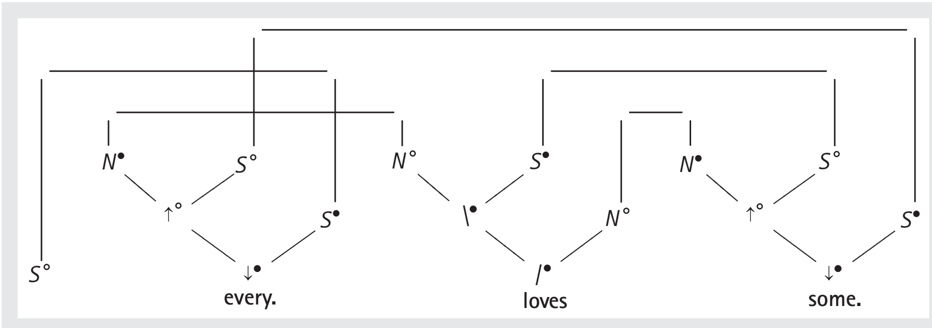


Figure 4.17. Proof net analysis for subject-wide scope *Everyone loves someone*

as the time-course of the load on working memory of a proof net analysis. For example, for the subject-wide scope analysis a of Figure 4.17 with semantics (29a) and the object-wide scope analysis b of Figure 4.18 with semantics (29b) the complexity profiles are as follows:

(30)

4		b		
3			ab	
2		a		
1	ab			
0				ab
	everyone	**loves**	**someone**	

The variation in the complexity profiles accords with the left-to-right quantifier scope preference.

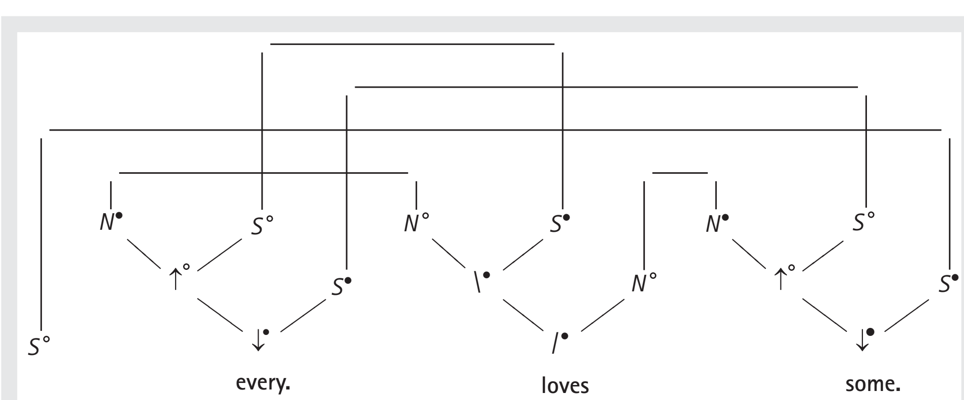


Figure 4.18. Proof net analysis for object-wide scope *Everyone loves someone*

CHAPTER 5

COGNITIVE GRAMMAR

RONALD W. LANGACKER

5.1 Contextualization

The central goal of Cognitive Grammar—and arguably of linguistics overall—is to describe the structure of particular languages and develop a general framework allowing the optimal description of any language. If this goal is widely shared, there is certainly no consensus about the proper means of achieving it. The chapters in this volume attest to the wide array of theories that grace the current linguistic scene. In no small measure, these alternative approaches stem from distinct conceptions of the target itself: they exhibit fundamental differences concerning not only the specific details of language structure but its very nature and status.

An initial question is whether the target exists at all. Is there an entity, reasonably called "the structure of a language", that is sufficiently coherent and well-delimited to be viable as an object of description? Indeed, given that no two people talk alike in all respects, is there something we can reasonably call "a language"? From the standpoint of Cognitive Grammar (CG), adopting either of two extreme positions on these issues would be equally misguided. One approach is to invoke an idealized speaker in a homogeneous speech community, ignoring such factors as acquisition, interaction, variation, registers, and language change. Now a certain amount of idealization is useful and unavoidable. When pushed to this extreme, however, it not only misportrays the nature of language but engenders an inaccurate account of its structure. The opposite extreme is to eschew idealization altogether.

The emphasis is thus on variation, in all its complex detail, to the point that expressions, rather than reflecting fixed linguistic structures, are seen as emerging in discourse through active negotiation by the interlocutors. While this does have an element of truth, its extreme version also misrepresents the nature of language, in effect denying that there is any structure to describe.

If properly conceived, the target does exist as a worthy object of description. There is, of course, no single, monolithic, or discretely bounded entity which can be identified as a language like English or Navajo. But at the same time, speech around the world is not a featureless sea of variation, with a purely random distribution of linguistic traits. To varying degrees, speakers are organized in cohesive groups whose members talk pretty much alike, at least compared with the speech of other communities. The groupings that emerge are real phenomena that merit investigation, just like waves, hurricanes, and galaxies. And as with such entities, a crucial factor is the basis for their emergence and cohesion—in this case linguistic similarities. Whatever their nature or extent, these are what we reify as the "structure" of a language or language variety. In practical terms, the description may have to focus on the most common variety, on traits that virtually all speakers share, or on the speech of a single individual who is somehow "representative". Though one is not then describing "a language" in its entirety, the task is essential even for purposes of studying variation. The latter can hardly be described or understood without describing particular varieties and knowing what it is that varies.

The "structure" of a language is not a discrete, static entity with a definite location (like the framework of a building). It is rather to be identified with certain patterns inherent in the interactive processing activity that constitutes speaking and understanding. Though intrinsically dynamic (as facets of processing activity), these patterns are often quite robust, with enough stability—through time and across speakers—to serve as the basis for negotiation and successful communication. They do not exist independently, as autonomous entities, any more than waves exist in isolation from an ocean. Instead these patterns, referred to in CG as *conventional linguistic units*, arise from other, more basic phenomena, consisting in particular ways of recruiting, adapting, and combining them. For example, phonological units (like segments, syllables, or words) represent specialized adaptations of more general capacities involved in hearing and producing sounds (e.g., pitch perception, pattern recognition, motor control, rhythm). Likewise, semantic units (the meanings of linguistic elements) draw upon the full range of conceptual abilities as well as any aspect of knowledge and experience. For analytic and descriptive purposes, we can hardly avoid discussing units as if they were fixed and separate entities. This reification should not however be projected onto the target. In the last analysis, units are best regarded as sociocultural skills: well-rehearsed patterns of processing activity called upon as needed as part of the more elaborate activity comprising linguistic interactions.

The descriptive goal of CG, as well as its view of the target, largely define its position within linguistics. Research on CG (originally called "space grammar") began in 1976, and its first accessible presentation was Langacker 1982. At the time it represented a radical alternative to standard linguistic doctrine, in particular generative grammar (Langacker 1987*b*: ch. 1), but, given how the field has evolved, today it is less of an outlier. CG shares with generative grammar the goal of explicitly describing language structure. However, it has a very different view regarding the nature of this structure and how it is properly characterized. CG belongs to the *functional* rather than the *formal* tradition, but its central claims make it quite distinct within the former.

The notorious division between functional and formal approaches is actually a gradation, ultimately a matter of whether functional considerations are viewed as being *foundational* or merely *ancillary* to the task of describing language structure (Langacker 1999*a*). Virtually all theorists would agree that language is shaped and constrained by the functions it serves: the *symbolic* function of pairing sounds with meanings; the *discourse* function of symbolic expressions; and the *communicative–interactive* function of discourse. It is further agreed that biological, psychological, developmental, historical, and sociocultural factors all have their place in a full account of language. At issue is the formalist contention that language structure is describable independently of such concerns as a prerequisite for addressing them. For functionalists, the autonomous entity thus described is an artifact: language does not exist, and cannot be properly characterized, independently of the factors cited. CG adopts the functional position, with the proviso that there are indeed structures to characterize. They arise as particular, conventionalized products of these factors, hence are not reducible to them, but have to be described explicitly.

Within the functional tradition, CG represents the movement known as *cognitive linguistics*. What merits the label "cognitive" is that, insofar as possible, language is seen as an integral facet of cognition, emerging from more general phenomena (e.g., perception, attention, categorization) rather than being separate and autonomous. The concern with cognition—not shared by all strands of functionalism—is fully compatible with the latter's emphasis on social interaction. It is only through interaction in a sociocultural context that language and cognition are able to develop. By the same token, an account of linguistic interaction cannot ignore the assessment by each interlocutor of the knowledge, intentions, and mental state of the other, as well as their apprehension of the overall discourse context. In its basic principles, CG (despite its name) strikes what is arguably a proper balance between cognitive and interactive factors (Langacker 2008).

CG is a particular cognitive linguistic theory. Even within cognitive linguistics, it stands out as radical due to certain basic claims, notably that grammar is wholly symbolic (hence meaningful) and that basic grammatical notions (like noun, verb, and subject) have unified conceptual characterizations. Nevertheless, it is largely compatible with a wide spectrum of cognitive and functional approaches, being general and flexible enough to incorporate their findings and insights. Research in CG per se has aimed primarily at developing a descriptive framework capable of handling the full range of structures encountered cross-linguistically; this is seen as just one aspect of ongoing cognitive-functional linguistic investigation concerned with prototypicality, typology, universals, and explanations for all the above. As a descriptive framework, CG most resembles Construction Grammar (Fillmore 1988; Goldberg 1995, 2006; Croft 2001; Fried and Boas 2005)—for instance, both view lexicon and grammar as a continuum consisting in *constructions* (form–meaning pairings)—but since the two developed in parallel and largely independently, there are also some basic differences (Langacker 2005*a*, 2005*b*).[1]

5.2 Interacting minds

Language is both cognitive and sociocultural. Language structure inheres in the interactive processing activity that occurs in discourse. It comprises a vast array of conventional linguistic units related to one another in various ways (e.g., by overlap, inclusion, or categorization). These units—established patterns of activity—reside in particular ways of recruiting and adapting other cognitive processes, with respect to which they are neither separate nor sharply delimited. This non-modularity has the consequence that the well-formed expressions of a language are not a well-defined set subject to algorithmic computation.

Discourse consists in series of *usage events*: instances of language use in all the complexity and detail evident to interlocutors. Depending on analytic purpose, a usage event can be of any size (e.g., a word, clause, sentence, turn, or "intonation unit" (Chafe 1998)). It subsumes both expression (audition, articulation, gesture) and full contextual understanding. The interlocutors function as *subjects* of conception, each apprehending an event from their own internal vantage point as a participant in it. The occurrence is nonetheless *intersubjective*, in that each participant apprehends the other and to some extent simulates their experience. It is through these interactive events that language structure is acquired, maintained, and modified.

Conventional linguistic units are abstracted from usage events through the reinforcement of recurring commonalities, at any level of specificity. Abstracted patterns are units by virtue of constituting well-rehearsed processing routines (*entrenchment*), and conventional by virtue of being shared by members of a speech community. Since both factors are matters of degree, there is no precise boundary between patterns that are and are not established in a language. The units that can

[1] Comparisons with some other frameworks are made in Langacker 1995*a* and 2004.

be posited are greatly restricted by the *content requirement*. They are limited, first, to elements necessary for language to fulfill its symbolic function: semantic units, phonological units, and symbolic units consisting in relationships between the two. Along another axis, they are limited to structures that are part of occurring expressions (hence directly apprehended), to schematizations of permitted structures, and to categorizing relationships between such structures. Thus units are either part of the primary data—usage experience—or derivable from it by means of basic psychological phenomena. Abstraction involves *schematization* because recurring commonalities are evident only at a certain level of granularity: usage events differ in their fine-grained details, which thus fail to be reinforced. The relation that abstracted units bear to usage events amounts to *categorization*. Categorizing relationships can themselves be reinforced and emerge as units. Precluded in this scheme are purely grammatical elements devoid of both conceptual and expressive content. By limiting structures to those directly grounded in usage, the content requirement affords an account of language acquisition which, at least in principle, is relatively straightforward.[2]

Structures reflecting any aspect of usage experience are capable of being abstracted as linguistic units. Most obvious are specific structures recurring as parts of expressions, e.g., the specific sound sequence [kæt], the concept [CAT], and the symbolic unit [[CAT]/[kæt]]. Among the schematic units derivable from these and comparable structures are the syllabic template [CVC], the concept [ANIMAL], and the symbolic structure [[ANIMAL]/[CVC]] (representing the commonality of nouns like *cat*, *dog*, *pig*, *sheep*, and *bear*). Established categorizations also count as units. Their status as such is essentially automatic when the target is fully compatible with the categorizing schema, as in [[ANIMAL]→[CAT]]. In this case the schema is *immanent* in the target, residing in aspects of the more elaborate processing activity constituting the finer-grained conception. Less automatic, hence more subject to conventional determination, are categorizations involving *extension* (not just *elaboration*), i.e., the target is in some way inconsistent with the categorizing structure. An example is the semantic extension [[[BEAR]/[bear]]- - →[[PANDA]/[bear]]], evident in the expression *panda bear*, wherein a panda is apprehended (at least for linguistic purposes) as a divergent type of bear.

Units like these represent only certain facets of usage events. Defined more broadly, usage experience unfolds in a number of *channels*, both *expressive* and *conceptual*. Segmental content is just one expressive channel, others being prosody and gesture. Likewise, conceptual channels are not limited to objective content (pertaining to the situation described), but also include information structure and speech management. Specifications in any channel or combination of channels can

[2] A promising basis for the account envisaged is the one proposed and empirically supported in Tomasello 2003.

be abstracted from usage events and incorporated in conventional linguistic units. In addition to sounds, for example, demonstratives often incorporate a pointing gesture. Old information is marked prosodically in English by reduction in pitch and stress. By way of speech management, we can signal the intent to keep talking ("hold the floor") either prosodically, through non-falling intontation, or segmentally, with the filler *uh* (*He's ... uh ... hard to please*).

An important part of experience in usage events is apprehension of the discourse they constitute. Thus many linguistic elements have meanings pertaining to the discourse itself (rather than the situation being described). For instance, words like *first*, *next*, and *finally* commonly indicate position in a discourse sequence, rather than a series of occurrences (*First, I'm too tired to discuss it. Next, I don't have time. Finally, I just don't want to*). We can reasonably say, moreover, that every element has some kind of discourse import. Even a noun like *cat*, which names a type of creature, carries with it the implication that this type needs to be specified in the discourse context; using the pronoun *one* instead (e.g., *this one*, in lieu of *this cat*) indicates the opposite. Because linguistic units are learned and maintained through usage, recurring aspects of their discourse function are themselves subject to entrenchment and conventionalization.

The same holds for other contextual factors. To the extent that an element consistently occurs in a particular kind of circumstance, schematic reference to that context will tend to be incorporated in the unit that emerges. This might involve some activity or occasion (e.g., saying *Cheers*! as a toast). It might involve properties of the interlocutors (e.g., *doggie* is only used with children) or their relative social status (as with honorifics). A form might be associated with a certain genre, register, or level of formality. It might be characteristic of a certain geographical region or social group. Through exposure to their use in context, a speaker thus gains either active or passive mastery of units representing numerous varieties of speech. The determining contextual factors are part of an expression's overall conceptual import, i.e., its meaning in the broadest sense. So in the CG account, specifications of a sociolinguistic nature do not require any special apparatus—they are simply abstracted from usage experience, like any other facet of conventional units. Indeed, elements which seemingly lack such specifications achieve their neutrality only through occurrence in a range of contexts sufficiently varied that specific contextual factors fail to be reinforced. But their decontextualization is never complete. As a minimum, every unit retains the generalized contextual specification of being used in speaking the language in question.[3]

Once abstracted from usage events, units are invoked in subsequent events. These established patterns of processing activity are essential resources exploited by interlocutors in speaking and understanding. Though not exhaustive of expressions, the

[3] Its status as part of a language is therefore one aspect of a unit's overall import. This is clearly relevant to the study of code-switching and multilingualism.

units activated in usage events provide the basis for their assembly, their structure, and their very interpretation as expressions of the language.

A linguistic expression only counts as such by virtue of being apprehended in relation to conventional units. Their relationship is one of categorization: particular units are invoked to categorize particular facets of a usage event. Apprehending even a small expression involves numerous categorizations pertaining to different facets of its conceptual, expressive, and symbolic organization. Collectively these constitute its *structural description*, i.e., its interpretation as an expression of the language. This relation of units to expression—referred to as *coding*—has to be effected by both interlocutors (in roughly commensurate fashion, if communication is to be successful). Coding is both individual and intersubjective. While the speaker engages primarily in *en*coding, and the hearer in *de*coding, to some extent each simulates the role of the other: the speaker assesses the expression's likely impact on the hearer, who anticipates what is coming.

Each facet of an expression's organization corresponds to any number of units with the potential to categorize it. But if a consistent interpretation is to emerge, only one can be activated for this purpose. How is it selected? Presumably the candidate units are mutually inhibitory and compete with one another for the privilege. Several factors then contribute to a particular unit becoming highly active, suppressing the others, and emerging as the winner. One such factor is a unit's degree of entrenchment, which translates into ease of activation. Another is the extent of a unit's overlap with the target, as assessed by preliminary, lower-level processing. A third factor is contextual priming, due to anticipation, recent occurrence, or the activation of associated structures. The end result is that the target is apprehended as an instance of a certain conventional pattern.[4]

Categorization does not imply full compatibility. Only as a special case is the categorizing unit wholly immanent in the target, so that their relationship is solely one of elaboration; in the rough and tumble of actual language use, we commonly push the limits by employing linguistic elements in ways not conventionally established. For instance, a word might be pronounced in a slightly divergent fashion, or understood with a somewhat different meaning. A categorization may then involve extension: the target is apprehended as an instance of the unit despite some inconsistency. When the disparities are sufficiently egregious, an expression is perceived as being non-conventional (i.e., as "ill-formed" or "ungrammatical"). It is fully conventional ("well-formed" or "grammatical") when all the categorizations in its structural description are elaborative.

A language is never static. The conventional units constituting a speaker's linguistic ability (or "knowledge" of a language) can only be maintained through usage. The reinforcement of units through their activation in usage events leads

[4] The target and categorizing unit should not be thought of as separate and distinct (recall that schemas are immanent in their instantiations). As with categorization in general, the unit—once activated—partially constitutes the target and shapes its apprehension in top-down fashion.

to their further entrenchment or at least prevents their decay. But usage also leads to their modification. Suppose a unit, [A], is used in a manner that conflicts with its specifications (e.g., with a slightly divergent meaning). This implies that it participates in a categorizing relationship of extension with (B), the target: ([A]- -→(B)). Should this happen on repeated occasions, both the target and its categorization as an instance of [A] are liable to be entrenched as units: [[A]- -→[B]]. And should this development become widespread in a speech community, these new structures achieve the status of conventional linguistic units, i.e., they are now part of the language. In this way usage engenders change. Moreover, since entrenchment and conventionality are matters of degree, there is never a sharp distinction between synchrony and diachrony. At a given point in time, the speech of an individual or a community always exhibits patterns with an intermediate status.

The adaptations induced by usage result in variation. Even for a single speaker, linguistic elements tend to have multiple variants: it is usual for a lexical item to have alternate senses (*polysemy*), for a phoneme to have alternate phonetic realizations, and for a construction to vary in form or conceptual import. Cognitive linguists thus recognize *complex categories*, each comprising an often substantial number of similar units linked by categorizing relationships (Lakoff 1987; Langacker 1987*b*: ch. 10; Taylor 2004). Normally category members center on a prototype, from which they develop radially by extension. As another dimension of category structure, members differ in their level of specificity, some constituting elaborations of other, more schematic units. To accommodate both dimensions, CG posits a network of related variants.[5] This is necessary since, for the most part, neither a prototype nor a high-level schema (should one be abstracted) accounts by itself for a speaker's full mastery of established patterns. Indeed, low-level regularities are often more important than global patterns for purposes of assembling expressions and assessing their conventionality. The essential role accorded to lower-level structures and usage-induced variation makes CG a *usage-based approach* (Barlow and Kemmer 2000; Langacker 2000; Bybee and Hopper 2001).

5.3 Conceptual semantics

How linguists think about grammar is greatly influenced by how they think about meaning. Approaches to meaning that bypass the role of human conception—treating it in terms of formal logic, truth conditions, or correspondences to the world—resonate with the view of grammar as an autonomous formal system.

[5] Though useful, the "network" metaphor may be overly discrete. An alternative is considered in Langacker 2006.

By contrast, meaning is taken in CG (and quite generally in cognitive linguistics) as residing in conceptualization.[6] This makes possible a symbolic account in which all grammatical elements have semantic import.

The term "conceptualization" is employed (in preference to "concepts") to emphasize the *dynamic* nature of linguistic meaning. It is dynamic in several respects. First, it includes not only fixed concepts but also the new conceptions forged in usage as the meanings of complex expressions. The negotiability of meanings, and hence their modification in context and through time, is a further aspect of their dynamicity. Another is the recognition in CG that linguistic structures—even established units—consist in processing activity. While it may seem instantaneous, apprehending the meaning of even a minimal expression (e.g., a lexical item) requires a certain span of processing time, however brief. Finally, the time course of conception (how it unfolds through processing time) is itself a dimension of linguistic meaning and thus a basis for semantic distinctions. Consider these expressions: *The scar runs from his elbow to his wrist* vs. *The scar runs from his wrist to his elbow*. They are not semantically equivalent despite describing the same objective situation and invoking the same conceived elements. The difference resides in sequence of mental access, i.e., the path followed in scanning along the scar to arrive at a full conception of the overall configuration.

Conceptualization is further understood as being *engaged*, *embodied*, and *interactive*. In addition to purely mental ruminations, it is taken as including sensory, motor, and emotive experience. Rather than being insular, therefore, conceptualization encompasses these basic modes of engaging the world. Experience at this level can hardly be divorced from the body and how it functions. Under the rubric of *embodiment*, cognitive linguists hold that this is true for cognition in general (Johnson 1987; Lakoff and Johnson 1999; Ziemke et al. 2007). We identify the brain as the primary locus of the processing activity comprising it. But the brain does not labor in isolation—it is embedded in the body, which in turn is embedded in the world. Cognition is grounded in the brain's functioning as an integral part of these larger systems. In particular, such grounding proves essential even for abstract conceptions related only indirectly to immediate physical experience, notably by furnishing the vehicle for metaphor (e.g., in expressions like *grasp the concept* or *see through his transparent lies*).

Equating conceptualization with the processing activity of individuals does not belie its interactive character. A crucial feature of the world we engage through conception is the existence of other conceptualizing individuals. We are adept at reading their intentions, simulating their experience, and assessing what they know. It is only through social interaction in a cultural context that language can be acquired

[6] Levinson (1997) wrongly ascribes to CG the position that conceptual and semantic representations are indistinguishable. Conceptualization only counts as linguistic meaning by virtue of being exploited and adapted for language purposes, hence shaped by convention and the functions served.

and cognition can fully develop. Furthermore, a large proportion of what counts as "the world" for conceptual and linguistic purposes is mentally and socioculturally constructed. Physical reality is only one facet of what we think and talk about (some others being money, religion, politics, social relationships, vacation plans, and so on endlessly). But despite this interactive basis, conceptualization is still carried out by individuals. Even when meanings are negotiated and co-constructed by the interlocutors, each apprehends the result as well as the nature of their respective contributions.

The conceptualization constituting linguistic meaning cannot be sharply delimited. This is so even for fixed expressions, i.e., lexical items, whose meanings are conventionally established. We have already noted some sources of indeterminacy in lexical meaning: negotiability (which can lead to semantic extension), polysemy (which results from extension), and the graded nature of entrenchment and conventionality. But even if we focus on a single, well-established sense, it cannot be characterized as a distinct conceptual entity of determinate scope. Consider the lexical unit *cat*, in its ordinary sense of referring to a typical domestic feline. Speakers know a great deal about such creatures: their size and shape; the range of their colors and markings; the sounds they make; how they move; what they eat; that they are kept as pets; their penchant for sleeping; that they like to scratch on furniture; their playfulness as kittens; their occasional utility in catching rodents; their stereotype as being self-centered and aloof; their cultural role in being emblematic of the mysterious; and so on. In contrast to standard doctrine, there is no good reason to believe that any particular portion of this knowledge functions exclusively as the linguistic meaning of *cat*, or that it assumes any special, specifically linguistic format. From the CG standpoint, the lexeme is seen instead as flexibly invoking this largely independent knowledge in context-dependent fashion.

Despite being flexible and open-ended, lexical meanings have the status of conventional units by virtue of representing established ways of evoking and viewing conceptual content. A unit of this sort provides mental access to a particular body of knowledge (e.g., everything associated with a certain type of creature). The access afforded is also conventionalized in that components of this knowledge vary in their baseline likelihood of activation: some specifications are highly central, so that their activation is essentially automatic; others are intermediate; and some are so peripheral as to be activated only in special contexts. Hence the conventional determination of lexical meaning is only partial. No doubt a lexeme has slightly different semantic import—as assessed by which specifications are activated, and to what degree—on every occasion of its use. It is nonetheless a conventional unit, and even non-central specifications are exploited for linguistic purposes. For example, the expression *cat-proof*, depending on whether it is applied to a birdcage, a sofa, or a rug, is interpreted with respect to different kinds of stereotypical cat behavior (catching birds, scratching on furniture, or excretion).

Lexical meaning is therefore viewed in CG as being fuzzily delimited, non-distinct from other knowledge, and only partially determined conventionally. This characterization is perhaps more evident when we turn to the composite meanings of complex non-lexical expressions. Here standard doctrine assumes a definite boundary between *semantics* and *pragmatics* (linguistic vs. extra-linguistic meaning) and takes the former as exhibiting *full compositionality*: the meaning of the whole derives in regular fashion from the meanings of the parts, as specified by rules of semantic composition. One problem with this scheme is the presumption that the parts (lexical items) have fully determinate meanings. Another is the absence of any specific line of demarcation between semantics and pragmatics. And while there are indeed patterns of semantic composition,[7] other factors come into play, so that semantics exhibits only *partial compositionality*.

We think of a sentence as having a coherent linguistic meaning based on the words it contains and their combination. Literally, however, an expression's meaning is not to be found "in its words", which merely serve as prompts for an often elaborate process of meaning construction (Reddy 1979; Fauconnier 1985; Levinson 1997; Langacker 2008). Its content, coherence, and overall import are commonly dependent on covert factors that are not exclusively or even primarily linguistic. One such factor is conceptual metaphor, residing in mappings between two domains of experience (Lakoff and Johnson 1980). For instance, the following sentence makes no literal sense, but we readily understand it by mapping our knowledge of ships onto the domain of theories: *His theory hasn't sunk yet, but it's taking on water and starting to list*. An expression's meaning may be crucially dependent on context or implicit background knowledge. In buying a digital camera, for instance, the specific meaning of *memory stick*—though not derivable from the individual meanings of *memory* and *stick*—is easily inferred when the term is first encountered. The background tacitly invoked can also be the product of imagination. Consider the apparent use of frequency adverbs (*always*, *often*, *seldom*, etc.) to quantify nouns: *Linguistic theorists are very often arrogant*. On the relevant interpretation, the sentence does not mean that any given theorist is frequently arrogant, but rather that *many* theorists are arrogant by nature. This meaning hinges on an imagined scenario of progressing through life, encountering linguistic theorists, and assessing them for arrogance. What *very often* specifies is the frequency of encounters in which the linguist in question exhibits this property.

So if we want to say that expressions "have" meanings, with enough substance and coherence to be recognized as such, they cannot be obtained from traditionally accepted linguistic elements (lexical meanings and patterns of semantic composition) working in isolation. Many other resources are exploited for this purpose, including a vast *conceptual substrate* and varied *conceptual abilities*. The former

[7] In CG, these are inherent in grammar rather than constituting a separate "semantic component".

provides the *content* of conceptions, while the latter reside in operations providing alternate ways of *construing* that content.[8]

An expression's content consists in a set of cognitive *domains*. Certain domains are said to be *basic*: space, time, and those associated with the senses (e.g., the range of colors we can perceive). These are not themselves concepts but irreducible realms of experience within which conception can emerge. Conceptions of any sort—regardless of area, complexity, or degree of entrenchment—qualify as *non-basic* domains. These run the gamut from minimal concepts (e.g., "yellow"), to those of modest complexity (e.g., "uncle"), to elaborate systems of knowledge (e.g., the rules and strategies of chess). They thus include, for example, all the notions cited above in relation to *cat*, as well as metaphor and imagined scenarios. Further included (since novel conceptions qualify) are apprehension of the context and the ongoing discourse. The domains invoked by a given expression are therefore numerous and diverse. Rather than being separate or disjoint, they are related in various ways (e.g., by inclusion, overlap, metaphorical mappings). What to identify as domains, and thus how many there are, is therefore somewhat arbitrary. The main point is simply that an expression's conceptual content is complex, multifaceted, and not limited to the domains evoked by its component lexical items.

Conceptualization is processing activity in which a conceptualizer (the *subject* of conception) apprehends some array of content (the *object* of conception). This content—even when pertaining to the immediate physical circumstances—is not just passively registered but apprehended through an active process of mental construction. Our conceptual abilities render the outcome non-determinate: the same content can be construed in alternate ways, resulting in different linguistic meanings.

The interlocutors function as primary conceptualizers for the meanings of expressions. Hence they figure in the meaning of every expression, if only in their minimal role as subjects of conception. In this capacity they remain implicit, since they are not part of the situation described. They nonetheless contribute to the expression's semantic import by virtue of apprehending its content and imposing a particular construal. To the extent that the interlocutors function only as subjects of conception, they are said to be *subjectively construed*. Generally, however, they also function to some extent as objects of conception, being part of the situation described, and, to that extent, in that capacity, they are said to be *objectively construed*. If nothing else, the interlocutors and the speech event tacitly define the deictic center. Even when left implicit, they are often invoked as participants in the coded situation (e.g., *He lives quite far away* [from us]; *That's impressive* [by my standards]; *She was very helpful* [to me]). But they can also be made explicit by personal pronouns (*I*, *you*, *we*), in which case they are an integral part of the expression's objective content.

[8] In a dynamic view, these two aspects of conceptualization cannot be sharply distinguished.

The subjects of conception adopt a particular *perspective* in "viewing" the objective content.[9] Most obviously, they assume a certain *vantage point*, by default identified with the deictic center. In *We'll go there tomorrow*, for example, the speech event functions as the spatial vantage point for *go* and *there*, as well as the temporal vantage point for *tomorrow*.[10] The vantage point adopted need not be the speaker's actual one and does not have to be consistent throughout; the speaker might also say *We'll come there tomorrow*, where *come*—in contrast to *there* and *tomorrow*—describes the journey from the vantage of its endpoint. Well known from discourse studies, the interplay of multiple vantage points, both actual and imagined, is thus characteristic of even simple expressions. Indeed, it figures in the meaning of the pronoun *I*, where the speaker functions as both the subject and an object of conception: the latter involves a simulation of how the speaker appears to other observers (Langacker 2007).

Other construal phenomena effect the *selection* of content for presentation. One is selection for *specificity*: the degree of precision and detail at which a situation is characterized. Through lexical and grammatical means, we can offer either schematic or progressively finer-grained descriptions (e.g., *thing* → *creature* → *animal* → *feline* → *cat* → *calico cat* → *lazy calico cat*). There is also selection for *scope*: the extent of the content an expression invokes as the basis for its meaning (its coverage in active domains). A distinction can often be made between an expression's overall conceptual content, or *maximal scope*, and the portion relied on most directly, its *immediate scope*. As its maximal scope, for instance, *elbow* invokes the conception of the body, but within this the arm is clearly most relevant and is thus its immediate scope. The maximal scope for *tomorrow* includes the speech event, which serves as temporal vantage point, and the structuring of time as a series of days. Its immediate scope is a limited span of time accessible from the vantage point when adopting a future *orientation* (another perspectival factor).

The selection of content represents an initial phase in the directing and focusing of attention. Through focusing for different purposes, at successive levels of organization, content elements are accorded various kinds and degrees of conceptual prominence. Special linguistic importance attaches to focusing done for referential purposes. Within its immediate scope, an expression singles out some particular element as the one it designates (refers to). This is called its *profile*. The profile of *elbow* is the part of an arm that bends, at its center. When used as a noun, *tomorrow* profiles the day following the speech event (not, say, some other day or time in general). Whereas an expression's immediate scope is the *general locus* of

[9] Certain aspects of visual perception have abstract analogs in general conception, hence the term "viewing". For the same reason Talmy (1996) speaks of "ception". However, CG does *not* claim that all conception is visuospatial in origin.

[10] In an earlier speech event, involving different interlocutors, the same projected journey might be described as follows: *They'll come here next week*.

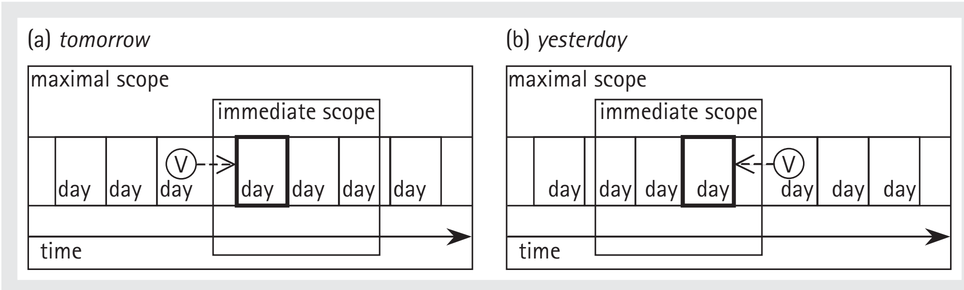


Figure 5.1.

viewing attention (described metaphorically as the "onstage" region), its profile is the *specific focus* of attention.[11]

Tomorrow is sketched in Figure 5.1(a), where V is the offstage vantage point, a dashed arrow represents orientation, and heavy lines indicate profiling. Though implicit and subjectively construed, V functions as point of access to the objective content and thereby serves to identify the profiled entity: starting from V, *tomorrow* designates the first day encountered in scanning toward the future. Comparison with *yesterday*, in 1(b), illustrates the basic point that expressions with the same conceptual content can nonetheless differ in meaning owing to the alternate construals they impose on it. The two words invoke the same conceptual elements. The semantic contrast—a matter of perspective and prominence—consists in orientation of the path of access from vantage point to profile.

An expression profiles either a *thing* or a *relationship*, each notion being defined abstractly in terms of mental operations.[12] For instance, *before* and *after* profile relationships in time, whereby one participant is situated relative to another. In Figure 5.2, a dashed arrow represents the profiled relation of temporal precedence, whose participants (given as boxes) are either events or temporal locations (e.g., *She left before noon*; *the day after our wedding*). Observe that the two expressions invoke the same content and have the same profile: in terms of what they refer to, a *before* relation is also an *after* relation. The semantic contrast resides in another level of focusing. Within a profiled relationship, one participant stands out as the *primary focus* of attention. Called the *trajector* (tr), this participant is the one the expression is construed as locating, describing, or characterizing. Often another participant stands out as a *secondary focus* by virtue of being invoked as a *landmark* (lm) for this purpose. The difference, then, is that *before* locates its trajector in relation to a later occurrence, and *after* in relation to an earlier one.

[11] In contrast to the subject of conception, the profile is salient, onstage, and explicitly mentioned, hence construed with maximal objectivity.

[12] See Langacker 2008: ch. 4. Here it is merely noted that things are not limited to physical entities and that the elements participating in a relationship need not be distinct, salient, or individually recognized.

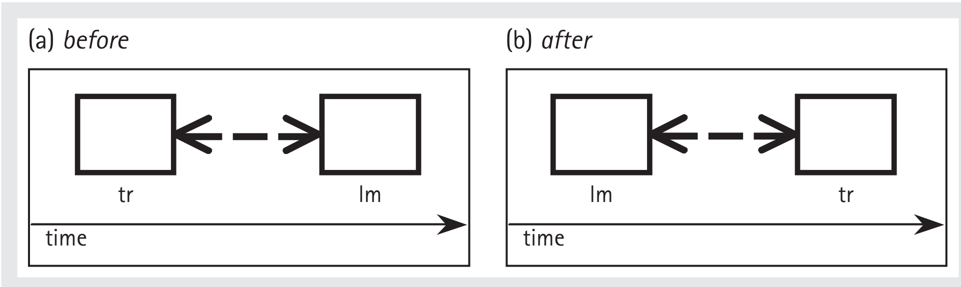


Figure 5.2.

Construal phenomena demonstrate that conceptualization is an active process of mental construction, both grounded in bodily experience and transcendent with respect to it. Its transcendence is due in no small measure to certain cognitive abilities inherent in immediate experience coming to be used independently. An example is our capacity for visual scanning, the dynamic process of directing attention along a path, as in following the flight of a bird or reading a line of text. The visual image involved in remembering or simply imagining the flight of a bird consists in certain aspects of this processing activity occurring independently, as a partial simulation of the scanning experience. This scanning is further disengaged from immediate perception in cases of "fictive motion" (Talmy 1996; Matsumoto 1997; Matlock 2004), e.g., *The scar runs from his elbow to his wrist*, where attention is directed along the path-like extension of a static entity. Still further removed from perceptual experience is the use of simulated scanning to mentally access a set of entities not aligned along a spatial path, as in the following: *From the brightest student in the class to the dumbest, they all work hard*.

Such examples illustrate the fundamental and pervasive role of *imaginative* phenomena in cognition and conceptual semantics. Fictive motion is just one case of the widespread invocation of fictive or "virtual" entities, even for describing actual occurrences (Langacker 1999*c*).[13] Also crucial to cognition is metaphor, whereby abstract domains are understood in terms of more concrete experience. Metaphorical mappings are a source of *blending*, in which selected elements from different domains are integrated to form a distinct conception with its own organization and emergent structure (Fauconnier and Turner 2002). As hybrid entities, such imaginative creations may be internally inconsistent—e.g., the ideas we *grasp* or *kick around* are conceived as being both physical and mental—and often cannot be "real" in any narrow sense. They are, however, an integral part of the mentally constructed world we experience and talk about, hence "real" in the sense that counts

[13] Suppose, for instance, that *Each boy threw a rock* describes an actual complex event. The referents of *each boy* and *a rock* are nonetheless virtual entities "conjured up" in order to make a generalization, not actual individuals. Likewise for the profiled event of throwing.

for linguistic purposes. It is largely through imaginative capabilities that, starting from bodily experience, our mental world is able to emerge in all its richness and complexity.[14]

5.4 Symbolic grammar

A standard approach to grammar, long predominant in modern linguistic theory, views it as being distinct from both lexicon and semantics. It is thus describable by purely formal combinatory rules based on a set of special, specifically grammatical primitives. Not itself meaningful, grammatical structure is one input to rules of semantic composition, which derive the meaning of a sentence from the meanings of the lexical items it contains.

CG is antithetical to this view in all respects.[15] Instead of being dichotomous, lexicon and grammar form a continuum of meaningful structures. Rather than being special and irreducible, therefore, the basic elements of grammatical description have both conceptual and expressive value. Lexicon and grammar reside in *assemblies of symbolic structures*, the main difference being that "grammatical" elements are generally more schematic. Being symbolic in nature, grammar is not distinct from semantics but incorporates it: semantic structure comprises the conceptual aspect of symbolic assemblies, their *semantic pole* (the expressive aspect constitutes their *phonological pole*). In particular, patterns of semantic composition are simply the semantic pole of *constructional schemas*—assemblies describing the formation of symbolically complex expressions. But despite the existence and importance of these patterns, linguistic meaning is only partially compositional.

It is not claimed that grammar is predictable from meaning.[16] The claim, rather, is that only meaningful elements are required for its proper description. Posing a major challenge in this regard are grammatical categories: the ad hoc classes defined by occurrence in particular constructions, as well as arguably universal categories like noun, verb, and adjective. The former are dealt with in a usage-based approach by representing constructions as complex categories, where high-level schemas capturing broad generalizations coexist with lower-level structures spelling out the details of their implementation. For example, the fact that *teach* participates in a certain minor pattern of past-tense formation is

[14] For their constitutive role in mathematics, philosophy, and culture, see Lakoff and Núñez 2000, Lakoff and Johnson 1999, Kövecses 2005.

[15] Only the briefest sketch can be offered here of the descriptive framework and its application. For a fuller picture, see Langacker 1987*b*, 1990, 1991, 1999*b*, 2008 and the references cited therein.

[16] It cannot be, since grammar itself contributes to an expression's meaning.

not specified by marking it with a meaningless diacritic but rather by including *taught* as one of the conventional units instantiating the schema describing the pattern. As for classes like noun and verb, standard dogma holds that they are not semantically definable. The usual arguments for this doctrine rest, however, on erroneous assumptions. For one thing, they presuppose an objectivist semantics, completely ignoring construal and imaginative phenomena. Moreover, they only consider characterizations at an intermediate level of abstraction—notions like "object", "action", and "property"—which are only suitable as category prototypes. Plausible definitions valid for all members have to be more schematic.

CG proposes characterizations based on cognitive abilities rather than any specific conceptual content. Suggested for nouns are two abilities quite apparent from collective nouns (*stack*, *herd*, *orchestra*, etc.): the grouping of constitutive entities, and their conception as a single entity for higher-level purposes. A noun designates a *thing*, defined as any product of these reifying operations.[17] Suggested for verbs are two abilities clearly evident in the real-time observation of events: conceiving of entities in relation to one another, and tracking a relationship in its evolution through time. A verb designates a *process*, defined as a relationship scanned sequentially in this fashion. Being maximally schematic in regard to content, these characterizations are plausibly ascribed to all class members. The extension of these categories to non-prototypical instances is a matter of cognitive abilities inherent in the prototypes (physical objects and events) coming to be used independently. And since these abilities are very basic, it stands to reason that the noun and verb categories are fundamental and universal (Langacker 2005*a*).

An essential point is that an expression's grammatical category depends specifically on the nature of its *profile*, not its overall content. In baseball, for example, the term *bat* designates either a long, thin wooden object, or else the action of using such an object in attempting to strike a ball. In the first case it profiles a thing, and is thus a noun, while in the second it profiles a process, which makes it a verb. Expressions that profile the same conceived entity may nonetheless belong to different categories by virtue of construing it differently. We can describe the same event, for instance, with either the verb *explode* or the derived noun *explosion*. Such examples are often cited as indicating the semantic non-definability of grammatical classes. Actually, though, they demonstrate the conceptual nature of linguistic meaning. Whereas the verb construes the event in unmarked fashion, as a process, the noun construes it as an abstract thing produced by conceptual reification.[18]

17 Usually (as with physical objects) these operations occur automatically at a low level of processing, so we are only aware of their product. They are made evident by their application at higher levels of organization to entities apprehended individually.

18 Its constitutive entities can be identified as the temporal phases ("time-slices") of the processual relationship (Langacker 1991: §1.2).

The descriptive notions of CG make possible the semantic characterization of other traditional categories. These are not however adopted exclusively or uncritically. Indeed, the conceptual descriptions reveal their limitations and lead to other categorizations. One class not traditionally recognized comprises expressions that resemble verbs because they profile relationships but resemble nouns by being non-processual. The profiled relation can fail to qualify as a process for several reasons. It may be non-processual because it is fully instantiated at a single point in time, as with adjectives and prepositions: while a person is usually *tall* or *in the kitchen* through some span of time, the relationship is fully manifested at each instant during that period. A stative participle, such as *broken*, is based on a verb but is non-processual because it profiles only the final, resultant state of the verbal process. An infinitival phrase, e.g., *to break*, retains the entire verbal relationship as its profile but renders it non-processual by viewing it holistically (typically with future orientation) rather than scanning it sequentially. As evidence that these elements do have something in common, observe that they can all be used to modify nouns (whereas verbs cannot): *a tall girl*; *the girl in the kitchen*; *that broken window*; *the first window to break*; **a break window*.

Relational expressions can be categorized based on the number and nature of their focal participants. For example, the basic difference between an adjective and a preposition is that the latter has a landmark as well as a trajector. Their content can be the same, as shown by pairs like *a nearby village* and *a village near the river*. Both indicate the proximity of their trajector (specified by *village*) to another entity. But while the preposition puts this onstage as a focused element (*the river*), the adjective leaves it implicit by virtue of being contextually identified. Another difference is that the trajector of an adjective is always a thing,[19] whereas a preposition is neutral in this regard. Thus a prepositional phrase can modify either a noun, in which case it functions like an adjective, or a relationship, in which case it functions adverbially (e.g., *They played near the river*). These brief observations illustrate a general point: although the traditional parts of speech are far from optimal (especially when taken as disjoint categories), and cannot be defined in terms of semantic content, the reason they have long proved useful for describing grammar is that they do have a conceptual basis, residing in construal.

The factors involved provide straightforward accounts of how expressions with roughly equivalent conceptual content can still be very different in their grammar. For instance, we can describe two entities as being similar with the verb *resemble*, the preposition *like*, or the adjective *alike*. These are sketched in Figure 5.3. Each profiles a relationship (given as a double-headed arrow) in which two things (shown as circles) come close to matching in some respect. The difference between the verb and the preposition is that the verb specifically portrays this relationship as

[19] Since they are non-processual and their trajector is a thing, participles and infinitives also qualify as adjectives and can modify nouns, as noted above.

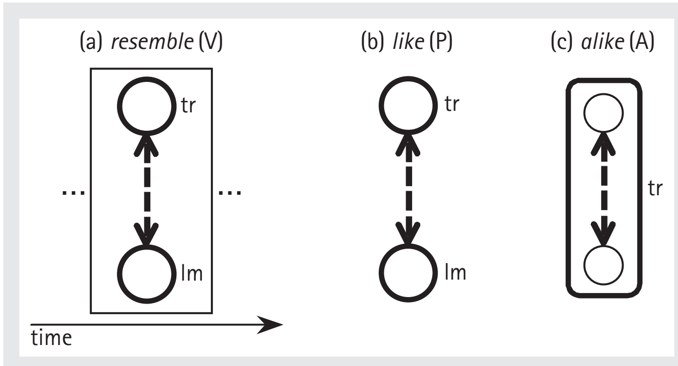


Figure 5.3.

continuing through time, making it processual, while the preposition abstracts away from time and designates the relationship as manifested at any given instant. In each case one participant is focused as trajector, i.e., as the one being assessed by comparing it to the other. Apart from their trajectors, *alike* is the same as *like*. The adjective does not confer focal prominence on either of the things compared, taken individually, but rather on the collective entity they constitute at a higher level of conceptual organization. The profiled relationship is thus symmetrical, as it holds between subparts of its single focused participant.

Since the crucial content is conveyed by members of different categories, clauses based on these elements are grammatically dissimilar despite their rough equivalence: *X resembles Y*; *X is like Y*; *X and Y are alike*. A finite clause profiles a process. As a verb, *resemble* itself meets this requirement and can therefore function as clausal head. The subject and object nominals then serve to specify its trajector and landmark.[20] On the other hand, *like* and *alike* are non-processual, so they cannot themselves head a clause. For clausal use they require the support of *be*, a highly schematic verb which merely indicates the continuation through time of some relationship. *Be* derives the composite processual expressions *be like* and *be alike*, which profile the continuation through time of the specific relationships designated by the preposition and the adjective. The trajector and landmark of *be like* are specified by nominals, just as for *resemble*. By contrast, *be alike* has only a trajector—a higher-order thing consisting of multiple individuals—hence it takes just a single nominal complement, either conjoined or plural. The resulting clauses are semantically equivalent in the sense of describing the same situation and having the same conceptual content. Their grammatical differences are a matter of using alternate symbolic means to arrive at this conception. Their shared content

[20] Also essential for a finite clause is *grounding* (Brisard 2002), effected in English by tense and the modals. Grounding (effected by the determiner system) is also characteristic of a *nominal* (or "noun phrase"). In both cases it establishes a basic connection between the profiled thing or process and the *ground*, i.e., the speech event and its participants.

notwithstanding, the expressions differ in meaning owing to the construal imposed by symbolic elements at each step in their composition.

The primary difference between "lexical" and "grammatical" units is that the latter are more schematic in their content, their main import residing in construal. On this basis *resemble* is considered a lexical verb, and *be* a grammatical element, even though each profiles an imperfective process.[21] As with many grammatical elements, *be* is the maximally schematic member of its class and serves a derivational function by imposing its construal on the specific content supplied by lexical items. *Be* is schematic for the class of imperfective verbs, since the relationship portrayed as continuing through time is otherwise unspecified. It imposes its processual construal on the specific relationship designated by an adjective or a preposition to derive a complex imperfective verb (e.g., *be alike*). Conversely, the stative participial morpheme imposes its non-processual construal on the content supplied by a verb, thus deriving an adjective (e.g., *broken*). A nominalizer (like the ending on *complainer* or *explosion*) is itself a schematic noun and derives a noun from the verb stem it combines with. The nominal profile is identified with a thing characterized in terms of the verbal process, such as a participant or the reified process itself.

Grammar consists in patterns for assembling simpler symbolic structures into more complex ones. At a given level of organization, *component* symbolic structures are *integrated*, both semantically and phonologically, to form a *composite* symbolic structure. These constitute an *assembly* of symbolic structures, which can either be specific or schematic. Specific assemblies represent expressions, either lexical (conventional units) or novel. Schematic assemblies—referred to as *constructional schemas*—are the "rules" of grammar. Constructional schemas are abstracted from symbolically complex expressions by reinforcement of their shared content and organizational properties. Once established as units, they can function as templates in forming new expressions and assessing their conventionality.

The symbolic structures constituting an assembly are linked by correspondences and relationships of categorization. Figure 5.4 shows this for the nominal expression *a woman like his mother* (determiners are omitted for simplicity). Consider first the lower level of organization, where the component structures *like* and *his mother* are integrated to form the composite structure *like his mother*. The preposition profiles a non-processual relationship (Figure 5.3(b)). The nominal profiles a thing (represented as a circle); M abbreviates its multitudinous semantic specifications (an open-ended set of cognitive domains). Effecting their semantic integration is a correspondence (dotted line) between the prepositional landmark and the nominal profile.[22] As a consequence, these two entities correspond to the same

[21] Processes are *perfective* or *imperfective* depending on whether the profiled relationship is specifically conceived as being bounded in time or as unbounded and internally homogeneous. The *count/mass* distinction for nouns is precisely analogous (Langacker 2008: ch. 5).

[22] This is symbolized by their phonological integration, wherein *like* directly precedes *his mother* in the temporal sequence.

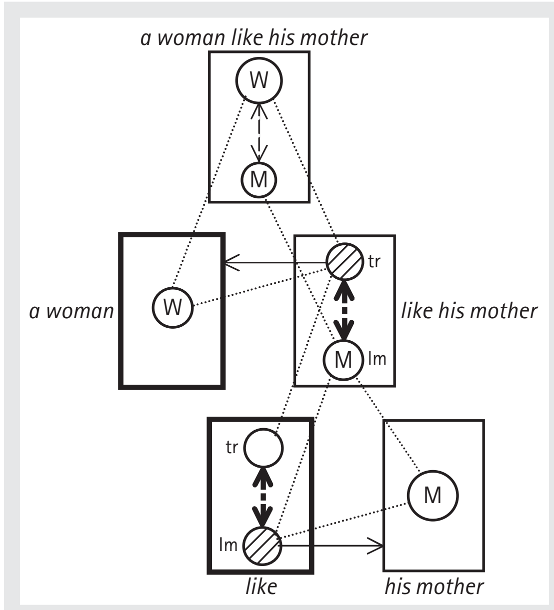


Figure 5.4.

entity at the composite-structure level, where the relational landmark merges their specifications. A solid arrow indicates that the nominal elaborates the preposition's schematic landmark, which is thus an *elaboration site* (marked by hatching). The heavy-line box indicates that *like* functions as the *profile determinant* in this construction, meaning that its profile is inherited as the composite-structure profile: *like* and *like his mother* profile the same relationship.[23]

At the higher level of organization, the component structures *a woman* and *like his mother* are integrated to form the composite structure *a woman like his mother*. Here the nominal elaborates the schematic trajector of the prepositional phrase, with the nominal serving as profile determinant. The overall expression is therefore nominal rather than relational, for it profiles a thing: an instance of *woman*, further characterized as resembling someone's mother.

A number of traditional grammatical notions can be characterized in terms of symbolic assemblies.[24] A nominal functions as a *subject* or an *object* by virtue of elaborating the trajector or the landmark of a profiled relationship (Langacker 1999*a*, 2008: ch. 11). A *head* is the profile determinant at a given level of

[23] In addition to the nominal elaborating the preposition's landmark, the component structures jointly categorize the composite structure. *Like his mother* constitutes an elaboration of *like* and an extension vis-à-vis *his mother*.

[24] As with the parts of speech, these notions are not accepted uncritically or as primitives, but are simply special cases of the configurations observed in symbolic assemblies. They are however quite typical, hence their continued descriptive utility.

organization: the component structure that imposes its profile—and thus its grammatical category—on the composite expression. *Like* is therefore the head within the prepositional phrase *like his mother*, and *a woman* within the nominal *a woman like his mother*. The distinction between a *complement* and a *modifier* depends on direction of elaboration between a head and the other component structure. A complement elaborates a salient substructure of the head, so in Figure 5.4 *his mother* is a complement of *like*. On the other hand, the head elaborates a salient substructure within a modifier, so *like his mother* functions as modifier with respect to *a woman*.

As seen in Figure 5.4, it is common for the composite structure at one level of organization to function in turn as component structure at a higher level. This is the basis for the traditional notion of *constituency*. To the extent that it emerges, constituency is real and readily accommodated in CG. It is not however considered essential or fundamental (more important are semantic factors like profiling, trajector/landmark alignment, and correspondences), hence it is often flexible, variable, and even indeterminate (Langacker 1997, 2008). In brief, constituency hierarchies are only one of the forms symbolic assemblies can assume, and a single hierarchy is seldom exhaustive of an expression's grammatical structure.

5.5 Justification

A succinct presentation of basic notions can merely hint at the nature and potential insight of CG descriptions. Its descriptive efficacy can in fact be cited as primary justification of the framework, which has been successfully applied to a wide range of phenomena in numerous and diverse languages. In particular, it has proved efficacious in dealing with a variety of classic problems—among them passives (Langacker 1982), "raising" constructions (Langacker 1995*b*), and pronominal anaphora (van Hoek 1995, 1997)—which have figured prominently in theoretical discussions framed by the autonomy thesis. It can be argued that the viability of CG's symbolic alternative to this thesis has been firmly established.

External justification for CG is considered quite important, and, while we must largely look to the future for such support, a fair amount is already available. First, a number of experimental findings can be cited in support of particular descriptive constructs. There is strong evidence, for example, that fictive motion (as in *The scar runs from his elbow to his wrist*) really does invoke the conception of motion for the construal of static situations (Matlock et al. 2005). Tomlin's research on attention in sentence processing (Tomlin 1995) gives evidence for the characterization of subject (trajector) as primary focal participant. More broadly, CG derives external support from its felicitous application in other language-related fields. For instance, it meshes quite well with Barsalou's model of cognitive processing (Barsalou 1999) as well as Tomasello's account of language acquisition (Tomasello 2003). It has implications for language pedagogy, where it has begun to be applied with some success (Pütz et al. 2001*a*, 2001*b*). It has further led to new initiatives in lexicography (e.g., Rivelis 2007).

A strong case can be made for CG being advantageous from the methodological standpoint. It may first be noted that a principled strategy has been followed quite consistently in the adoption of basic descriptive constructs—notions like profile, trajector, immediate scope, and sequential scanning. In each case they are posited on the basis of converging evidence of at least three kinds: psychological, semantic, and grammatical (Langacker 1993). Consider profiling. As the focusing of attention (for referential purposes), it meets the requirement of manifesting an evident psychological phenomenon. Next, it is shown to be important for semantic description. Making this quite apparent are expressions that invoke the same content but differ in meaning due to the alternate profiles they impose on it, e.g., *break* and *broken*. While all the content of *break* is also invoked by *broken*, they are semantically distinct because the former profiles the full event and the latter just the resultant state. Finally, having been established independently on semantic grounds, the construct is shown to play a role in grammar. Most notably, an expression's profile determines its grammatical category. *Break* is thus a verb, and *broken* a stative-adjectival participle, because they respectively profile a process and a non-processual relationship.

It can further be argued that CG has the methodological virtues of being quite restrictive in what it postulates while also achieving significant conceptual unification. Most broadly, it relies exclusively on cognitive phenomena that are either well known or easily demonstrated, thereby affording an explanation for many aspects of semantics and grammar. More specifically, the content requirement imposes severe restrictions on what can be posited: linguistic units are limited to semantic, phonological, and symbolic structures that are either part of occurring expressions or arise from them through abstraction and categorization.[25] A major source of conceptual unification is the characterization of lexicon, morphology, and syntax as a continuum consisting solely in assemblies of symbolic structures. CG employs the same basic descriptive apparatus for structures at any position along this spectrum. While these vary in specifics (e.g., symbolic complexity, degree of specificity, being fixed or novel), they all have both conceptual and expressive import. Nothing more need be posited. Rather than being separate and autonomous, grammar can be identified with certain facets of symbolic assemblies, which comprise just what is needed for language to serve its symbolic function of expressing meaning.

[25] Thereby precluded, for example, are derivations from underlying structures, as well as "purely grammatical" elements devoid of both semantic and phonological content.

CHAPTER 6

EMBODIED CONSTRUCTION GRAMMAR

JEROME FELDMAN
ELLEN DODGE
JOHN BRYANT

Recent developments in neuroscience and the behavioral sciences suggest approaching language as a cornerstone of Unified Cognitive Science. One such integrative effort has been under way for two decades in Berkeley. The NTL (Neural Theory of Language) project studies language learning and use as an embodied neural system using a wide range of analytic, experimental, and modeling techniques. The basic motivation for NTL and its relation to ongoing experimental work is discussed in several places.[1] The core idea is to take all the constraints seriously and to build explicit computational models that demonstrate the theoretical claims. At one level, NTL continues the tradition of Cognitive Linguistics (CL) represented by several chapters in this volume. But explicit computational modeling demands greater precision than is possible with the pictorial diagrams that remain standard in most CL work.

CL and related approaches to language stress the continuity of language with the whole mind and body and with society. Statistical considerations and

[1] The ECGweb wiki can be found at http://ecgweb.pbwiki.com/. A web search using ECG NTL ICSI will also work.

incremental learning and adaptation are also deemed essential parts of the capacity for language. The challenge is to develop a methodology that honors the inseparability of language use from cognition while being sufficiently rigorous to support formal and computational analysis. The NTL approach is to postulate distinct levels of description, explicitly mirroring the levels in the natural sciences such as biology, chemistry, and physics. The discussions in this chapter will focus on *computational level* descriptions of fairly complex language phenomena. In other work (Feldman 2006), we suggest how such descriptions can be reduced to a *structured connectionist level* and then to postulated brain structures. There is now a fair body of behavioral and biological experimentation exploring these models (Boroditsky 2000; Gallese an Lakoff 2005; Hauk et al. 2004).

Within the computational level, the NTL approach separates language understanding into two distinct phases called *analysis* and *enactment*. Schematically, analysis is a process that takes an utterance in context and produces the best-fitting intermediate structure called the Semantic Specification or *semspec* (cf. Figure 6.6). The semspec is intended to capture all of the semantic and pragmatic information that is derivable from the input and context. As we will see, this is at a rather deep level of embodied semantics. The semspec is used to drive inference through mental simulation or, as we call it, enactment. Within NTL, enactment is modeled using executing networks or *x-nets* which model the aspectual structure of events, and support dynamic inference (Narayanan 1999).

The grammar formalism of NTL is called Embodied Construction Grammar (ECG). It is a notation for describing language that is being used in a wide range of theoretical and applied projects. The ECG formalism is designed to be all of the following:

(1) A descriptive formalism for linguistic analysis
(2) A computational formalism for implementing and testing grammars
(3) A computational module for applied language tasks
(4) A cognitive description for reduction and consequent experiments
(5) A foundation for theories and models of language acquisition.

Embodied Construction Grammar is our evolving effort to define and build tools for supporting all five of these goals.

This chapter will focus on three related features of ECG: deep semantics, compositionality, and best-fit. The NTL theory behind ECG highlights two aspects of neural embodiment of language—deep semantics and neural computation. One can formalize deep semantics using ECG *schemas*. This includes ideas such as goals and containers, which have been at the core of Cognitive Linguistics from its origins (Lakoff 1987). As we will see, ECG schemas such as the EventDescriptor (Figure 6.3) can also describe much broader concepts. ECG schemas can also be used to represent the linguistically relevant parameters of actions (as we will see below), which are packaged in the semspec.

NTL posits that the key to language analysis lies in getting the conceptual primitives right, which in turn depends on evidence from biology, psychology, etc. As linguists, we evaluate putative primitives by their ability to capture general linguistic phenomena. ECG provides a mechanism for expressing and (with the best-fit analyzer) testing linguistic explanations. But isolated phenomena do not suffice for eliciting powerful primitives; we need to examine how a range of cases can be treated compositionally. The tools provided by the ECG system become increasingly important as the size of the grammar increases, as they facilitate testing a wide range of examples and thus help greatly in the cyclic process of hypothesizing linguistic primitives and using these to model complex phenomena.

The core of the chapter is a detailed analysis of a set of related constructions covering purposeful action, with an emphasis on compositionality. The examples illustrate the notation and central ideas of the ECG formalism, but hopefully will also convey the underlying motivations of NTL deep semantics and conceptual composition.

6.1 The position of NTL and ECG in the current study of language

The general NTL effort is independent of any particular grammar formalism, but it is strongly aligned with integrated approaches to language including several chapters in this handbook: Bybee & Beckner (this volume), Caffarel (this volume), Fillmore & Baker (this volume), Langacker (this volume), and Michaelis (this volume). Jackendoff (this volume) presents a different perspective, preserving the separation of form and meaning, but linking them more tightly than earlier generative theories.

NTL also suggests that the nature of human language and thought is heavily influenced by the neural circuitry that implements it. This manifests itself in the best-fit ECG constructional analyzer that is described later in this chapter. An important aspect of both NTL and of the ECG analyzer is the dependence on a quantitative best-fit computational model. This arises from the computational nature of the brain and shares this perspective with statistical (Bod, this volume) and Optimality (Gouskava; de Swart and Zwarts, this volume) approaches to grammar.

One way to characterize the ECG project is as formal cognitive linguistics. ECG is a grammar formalism, methodology, and implementation that is designed to further the exploration and application of an integrated, embodied approach to language. The explicit simulation semantics of NTL plays an important role in ECG because the output of an ECG analysis (cf. Figure 6.6) is a semspec.

At a technical level, ECG is a unification-based grammar, like HPSG (http://hpsg.stanford.edu/) and LFG (Asudeh, this volume) in which the mechanisms of unification and binding are extended to deep embodied semantics, discourse structure, and context, as we will show. A unique feature of the ECG notation is the **evokes** primitive, which formalizes Langacker's idea of a profile-base relation and models one aspect of spreading activation in the brain.

ECG is a kind of Construction Grammar (Goldberg 1995; Michaelis, this volume) because it takes as primitive explicit form–meaning pairs called *Construction*s. Both ECG schemas and constructions are organized in an inheritance lattice, similar to that described by Michaelis (this volume). ECG is called embodied because the meaning pole of a construction is expressed in terms of deep semantic schemas, based on postulated neural circuits and related to the image schemas of CL (Lakoff 1987). An explicit limitation is that no symbolic formalism, including ECG, can capture the spreading activation and contextual best-fit computations of the brain. In the conclusions, we will briefly discuss how NTL tries to unify ECG with neural reality.

6.2 Compositionality

Before introducing the technical details of the ECG formalism and illustrating their application in a more detailed case study, it may be helpful to first look at some of the challenges of compositional analysis and some ways these might be addressed.

One challenge is presented by the fact that a given verb will often exhibit more than one pattern of argument realization. For instance, *slide* appears in sentences such as: *The chair slid*; *Jack slid the chair*; and *Jack slid her the chair*. The verb's "slider" role is expressed in each of these sentences (*the chair*), but in each case is associated with a different grammatical role. Moreover, these sentences describe different types of events, differing as to the presence/absence of causation and transfer of possession, and express different numbers and types of semantic roles.

Using argument structure constructions (Michaelis, this volume), one can handle examples such as these without necessarily positing different verb senses for each pattern. The argument structure construction provides semantic roles associated with the basic type of scene being described, and specifies relations between these roles and grammatical arguments. Verbs have semantic roles associated with them, but do not have to specify how these are linked to grammatical arguments. A given verb can potentially combine with different argument structure constructions, each of which may describe a different type of scene. When a verb unifies with a specific argument structure construction, some or all of their semantic roles will unify.

The meanings and semantic roles associated with the verb and argument structure constructions are clearly a central component of such an approach, but the semantic representations commonly used are often inadequate in several respects. To fully support a compositional analysis of a broad range of examples, constructional meaning representations should meet the following criteria:

- Since it is not entirely predictable, the exact pattern of how the roles of each construction compose with one another should be explicitly specified. For instance, to analyze *The chair slid*, there needs to be a specification of the fact that the "sliding thing" verb role unifies with the argument structure construction's "Theme" role.
- Because a given argument structure construction will unify with many verbs, its specifications should not be lexically specific. We don't, for example, want to have to list every verb-specific role that an argument structure's "Theme" will unify with. Instead, more general specifications that capture the meaning common to these different verb roles are needed.
- Verb meaning should be represented so that it explicitly indicates the semantic similarities that motivate the unification of verb roles with the roles of argument structure constructions. This is not possible if verb meaning is defined too generally. For instance, we could use a "Theme" role to represent the sliding thing, motivating its composition with an argument structure construction's "Theme" role. But, this would not indicate what motivates it to compose with the "Patient" role in *He slid the chair*. On the other hand, if the argument structure construction for this second example is defined using a Theme role instead of a Patient role, the semantic distinction between transitive and intransitive constructions is obscured.
- The results of composition should be something more than just the conjunction of two role names (such as slider-patient or theme-patient); ideally, complex semantic roles should be defined using separately motivated conceptual structures.
- Meaning representations should capture both differences and similarities of meaning. Roles that are only defined at a very general level, such as thematic-type roles, support recognition of semantic similarity but obscure differences. More specific roles can make semantic differences more apparent, but need to be defined in such a way that it is also possible to recognize semantic similarities that motivate the composition of verb and argument structure constructions.

Each of these goals presents some challenges. Taken as a whole, it is clear that semantic representations need to include more than just very general thematic-type roles and very specific lexically-defined roles.

In addition to verb and argument structure constructions, a compositional analysis of sentence meaning requires that various phrase-level constructions be defined. To support a compositional analysis, these constructions need to be defined in such a way that they will unify with other constructions instantiated in a sentence, with the result that the composed meanings of these constructions specify the sentence's meaning. Therefore, phrasal constructions need to be coordinated with the definition of argument structure and other constructions.

What phenomena do such constructions need to deal with? In the case study presented later in this chapter, we will analyze a set of sentences which describe the same type of events and actions, but which exhibit other, sometimes subtle, differences in meaning. As a prelude, we will briefly describe some of the challenges presented by sentences such as these. To start, consider *She kicked the table, Her foot kicked the table*, and *The table was kicked*, sentences which can all be used to describe the same scene, but which differ in terms of scene perspective, or "participant profiling". While the same set of semantic roles are conceptually present in each case, there are differences as to how (and if) they are expressed. Consequently, such sentences may instantiate different, but related, argument structure constructions.

The sentences *You kicked the table, Did you kick the table?*, Which *table did you kick?, and Kick the table!* can all be used to describe the same type of kicking event. The kicking is in each case described from the same perspective (the kicker), which suggests that these sentences may all instantiate the same argument structure construction. However, they clearly differ as to their discourse purpose, as well as differing as to topic, presence/absence of auxiliary, and in other ways, suggesting the need for different phrase-level constructions specifying these differences. Ideally, the phrase-level constructions should be defined in such a way that they will unify with a range of argument structure constructions. So, for instance, they could also be used in the analysis of sentences such as *Did her foot kick the table? Which table was kicked?*, etc.

Ultimately, the goal of a compositional constructional analysis of an utterance is one of conceptual composition: when the constructions instantiated in a sentence unify, their meanings should compose, and this composition should represent the conceptual structure associated with the utterances as a whole. Consequently, it is critical not just to get the constructional "decomposition" of utterances right, but also their conceptual decomposition. But how can we get this right? How can we carve up conceptual structure along the proper "joints"?

Cognitive linguistics research has yielded many insights into the nature of the conceptual structures conveyed by language. Research on the meanings of spatial relation terms has shown that they can be analyzed in terms of combinations of primitive elements (e.g., Talmy 1972; 1983; Langacker 1976; 1987; this volume). These primitives include such things as bounded regions, paths, contact, etc. Cognitive linguists have also observed that primitives such as these recur across many different experiences, including but not limited to language. They have described various image schemas, each of which includes a relatively small numbers of parts or **roles**, and which have an internal structure that supports inferences (Johnson 1987). Later research supports the idea that such schemas are "embodied" neural structures (Regier 1996; Dodge and Lakoff 2005). Moreover, many types of culturally-specific experiences also exhibit schema-like structures, or frames (Fillmore 1982).

ECG builds upon these ideas, representing meaning using **schemas**, which are partial representations of neurally-based conceptual structures. Importantly, these schemas are defined independently of specific linguistic constructions. For example, we can define a Container schema, which can be used to represent the meaning of *in*. But, this schema's structure is also present in many non-linguistic experiences, such as experiences of putting things into and taking them out of various containers. Moreover, this same schema can be used in the meaning representations of many different constructions, such as those for *inside*, *inner*, *out*, and *outside*. Several examples are given in Figure 6.1, in a formal notation to be defined below.

Many complex conceptual structures can be "decomposed" and be represented as combinations of more "primitive" conceptual structures. For instance, going into a bounded region can be decomposed into a spatial relations component (a relation to a container) and a motion component. Crucially, the conceptual gestalt is a particular combination of these parts, not just some sort of random juxtaposition.

The meaning of many constructions can also be decomposed and represented using combinations of more primitive schemas. For example, the meaning of *into* can be represented as a unification of Container and SourcePathGoal schemas. And, when constructions compose, their meanings will also compose, generally resulting in more complex conceptual structure. Consequently, the meaning of the unified constructions can also be represented as a composition of schemas. Different constructional combinations will be associated with different complex conceptual structures, represented by different compositions of schemas.

To represent the meanings of a wide range of constructions, a relatively large, comprehensive inventory of schemas is needed. In addition to defining schema primitives that are used in descriptions of spatial relations, we also need schemas that represent other types of concepts, especially those that are directly relevant to argument realization, such as action, force, causation, motion, and change. But, it is not a simple matter to develop such an inventory. To do so requires that we figure out what sorts of more "primitive" conceptual structures might exist, as well as how other, more complex structures might be analyzed as involving productive compositions of these more primitive structures.

Our strategy is to examine a broad range of situations evidencing similar types of conceptual structures. By comparing them, we can gain insights into the types of distinctions and similarities that schemas representing such structure should capture. Further insights can be gained by using more than one kind of evidence to make these comparisons. We will briefly illustrate this strategy by examining some different types of motor-control actions, and discussing the sorts of schemas that seem to be needed to represent the conceptual structures associated with them. The following discussion uses linguistic examples to illustrate these comparisons, but there is also other cognitive and neural evidence to support these decompositions. A more detailed analysis is presented in Dodge (2009).

We start by considering motor control actions and the different types of "roles" they may include. For example, actions such as sneezing, walking, or smiling involve an actor and (typically) some part of the actor's body, but not necessarily another entity. And, verbs describing such actions typically occur in utterances which express this actor role, e.g., *He walked/sneezed/smiled*. Many actions also involve an additional entity, but the role of this entity differs in some important respects. Actions such as looking and pointing do not necessarily involve contact or transfer of force to this other entity. And verbs for these actions typically occur in utterances that express this other "target entity" role, but in such a way that does not indicate any physical affectedness, e.g., *He looked/pointed at her*. For actions such as kicking, squeezing, and pushing, however, the actor contacts and transfers force to another entity. In many cases, these forceful actions result in some physical change to the entity being acted upon, suggesting an important distinction in affectedness between forceful and non-forceful actions. These distinctions suggest the need for a basic motor-control schema with an actor role, and at least two other related schemas (for forceful and non-forceful actions) each of which contain some kind of role for an additional entity.

Next, let us look a little more closely at the relation between forceful actions and cause-effect. Forceful action verbs often occur in utterances that express the "acted upon" entity in a way that indicates it is affected in some way, e.g., *He squeezed/pulled/kicked the bottle*. But, the same action can potentially produce many different effects. For instance, kicking may make a leaf move, break a pane of glass, or cause someone pain. And, in some cases, it may not cause any perceptible effect, e.g., *He kicked at the door/pulled on the rope, but it wouldn't budge*. Furthermore, in many cases, any "effects" that do occur are ones that can also occur independent of the action (and, in some cases independent of any readily observable "causer"). For example, a leaf may fall off a tree, a window may break in a storm, and we may feel a sudden pain in our leg and not know what caused it. These observations suggest that schemas for forceful actions, as well as those for possible effects that are caused by such actions (such as motion and change of state), should be defined independent of causation. However, these same schemas can also serve as parts within more complex schemas, such as those for cause-effect events that involve a causal relation between a forceful action and the motion or change of some other entity. The remainder of this chapter formalizes these notions and shows how they can be combined with innovative computational tools to support deep semantic analysis of complex utterances.

6.3 ECG NOTATION AND PRIMITIVES

In ECG, construction grammars are specified using two basic primitives: **constructions** and **schemas**. Constructions are paired form constraints and meaning constraints. ECG is different from other construction grammar formalisms because the meaning constraints are defined in terms of embodied semantic schemas, such as those in Figure 6.1.

There are four ways to specify relations between ECG primitives: roles, subtyping (through the **subcase of** keyword), evoking a structure (through the **evokes** keyword), and constraints (co-indexation and typing). A **role** names a part of a structure, and the **subcase of** keyword relates the construction/schema to its type lattice, allowing for structure sharing through (partial) inheritance.

Evoking a structure makes it locally available without imposing a part-of or subtype relation between the evoking structure and the evoked structure. Using Langacker's standard example, the concept hypotenuse only makes sense in reference to a right triangle, but a hypotenuse is not a kind of a right triangle, nor is the right triangle a role of the hypotenuse. The evokes operator is used to state the relationship between the hypotenuse and its right triangle.

Like other unification-based formalisms, ECG also supports constraints on roles (features). The double-headed arrow operator is used for co-indexing roles (↔). Roles can be assigned an atomic value using the assignment operator (←). A type constraint (specified with a colon) constrains a role to only be filled by a certain type of filler.

The specific grammar described here will be called EJ1. Figure 6.1 shows a set of EJ1 semantic schemas ranging over conventional image schemas (TL and SPG), embodied processes (Process, ComplexProcess, and MotorControl), and motion schemas (Motion, TranslationalMotion and Effector TranslationalMotion). The TL schema has roles for a trajector and a landmark. The SPG schema inherits the trajector and landmark roles by subcasing TL, and adds roles for describing a path including source, path, and goal. As we will see in Figure 6.6, the embodied semantics (semspec) of ECG consists entirely of schemas with bindings between their roles.

The Process and ComplexProcess schemas are general descriptions of actions and events in which a single participant is profiled using the protagonist role. A ComplexProcess is made up of two sub-processes called process1 and process2. The ComplexProcess schema shows how roles can be bound (co-indexed)—required to have the same filler. The ComplexProcess's primary protagonist role (inherited from Process) is co-indexed (using ↔) with the protagonist role of process1, and the secondary protagonist2 role is bound to the protagonist of process2.

The Motion schema describes a process in which the mover is the protagonist and has a speed and heading. The TranslationalMotion schema is a subcase of Motion and

```
schema TL
  roles
    trajector
    landmark
```

```
schema SPG
  subcase of TL
  roles
    source
    path
    goal
```

```
schema Process
  roles
    x-net : X-Net
    protagonist
```

```
schema ComplexProcess
  subcase of Process
  roles
    process1: Process
    process2: Process
    protagonist2
  constraints
    protagonist ⟷ process1.protagonist
    protagonist2 ⟷ process2.protagonist
```

```
schema MotorControl
  subcase of Process
  roles
    actor
    effector
    effort
  constraints
    actor ⟷ protagonist
```

```
schema Motion
  subcase of Process
  roles
    speed
    heading
    mover
  constraints
    mover ⟷ protagonist
```

```
schema TranslationalMotion
  subcase of Motion
  evokes SPG as spg
  constraints
    mover ⟷ spg.trajector
```

```
schema EffectorTranslationalMotion
  subcase of ComplexProcess
  roles
    process1: MotorControl
    process2: TranslationalMotion
    target
  constraints
    process1.effector ⟷ process2.mover
    target ⟷ process2.spg.landmark
```

Figure 6.1. ECG representations of image schemas (TL and SPG), embodied processes (Process, ComplexProcess, and MotorControl), and motion (Motion, TranslationalMotion and EffectorTranslationalMotion)

adds the constraint that the motion is conceptualized as occurring along a path. The TranslationalMotion schema shows an example of the evokes keyword; the path is represented by the evoked SPG schema. In both schemas, the mover is defined as the protagonist using a co-indexation constraint. Each process has a role called x-net that further specifies the kind of action that is modeled by the schema. For example, the Motion process can describe walking, crawling, and many other methods of motion. Specific aspects of particular actions are represented by the filler of the x-net role. As a consequence, the Motion schema acts as an abstraction over all the different motion x-nets. This is one crucial requirement for compositionality.

The MotorControl schema has a special significance in the grammar. It is the semantic root of embodied, controlled processes. It adds roles for an actor, effector, and effort. The actor is the embodied protagonist, the effector is the controlled body part, and the effort is the energy expenditure.

The schema EffectorTranslationalMotion puts all of these schemas together to represent the idea of an entity controlling the motion of an effector. Effector-TranslationalMotion is a ComplexProcess in which process1 has a MotorControl type constraint and process2 is constrained to be a TranslationalMotion. The schema also adds a role for the target toward which the effector is moving.

Figure 6.2 shows the schemas that are combined to define the meaning of verbs of impact such as "hit", "slap", "kick", etc. The ForceTransfer schema describes a transfer of a particular amount of force between any kind of supplier and recipient. The ForceApplication schema describes MotorControl actions in which force is applied, and thus evokes the ForceTransfer schema. The ForceApplication schema adds roles for an instrument and an actedUpon entity. The constraints block of this schema then binds the appropriate roles. As we discussed earlier, a judicious choice of embodied schemas enables us to capture conceptual regularities and the ECG formalism supports this.

6.4 ECG CONSTRUCTIONS

Constructions are pairings of form and meaning, and, in ECG, this pairing is represented by a form block (defined by the **form** keyword) and a **meaning** block. Both the form pole and meaning pole of a construction can be typed. In this chapter, we simply constrain the form pole of HitPastTense and SlapPastTense to be a word using the Word schema (not shown). Form blocks can also have form constraints, and in these simple lexical constructions, the form constraint specifies that the orthography of the HitPastTense construction is "hit" by binding the string "hit" to self.f.orth. The slot chain self.f.orth uses the **self** keyword to refer to the construction itself, and the keyword **f** to refer to the construction's form pole. The Word schema has a role called orth to represent the orthography of a word.

Figure 6.2 also shows how past tense lexical constructions for "hit" and "slapped" can be defined in terms of the ForcefulMotionAction schema. Like schemas, constructions are arranged into a subcase lattice and the HitPastTense and SlapPastTense constructions subcase a general PastTense construction (not shown) that specifies facts about the tense and aspect of PastTense verbs.

```
schema ForceTransfer
  roles
    amount
    supplier
    recipient
```

```
schema ForceApplication
  subcase of MotorControl
  evokes ForceTransfer as ft
  roles
    instrument
    actedUpon
  constraints
    instrument  ⟷ effector
    effort  ⟷ ft.amount
    actor  ⟷ ft.supplier
    actedUpon  ⟷ ft.recipient
```

```
schema ForcefulMotionAction
  subcase of ComplexProcess
  roles
    process1: ForceApplication
    process2: EffectorTranslationalMotion
  constraints
    protagonist  ⟷ protagonist2
    process1.actedUpon  ⟷ process2.target
    process1.instrument  ⟷ process2.effector
```

```
construction HitPastTense
  subcase of PastTense
  form: Word
    constraints
      self.f.orth  ⟵  "hit"
  meaning: ForcefulMotionAction
    constraints
      self.m.x-net  ⟵  Hit-X-Net
```

```
construction SlapPastTense
  subcase of PastTense
  form: Word
    constraints
      self.f.orth  ⟵  "slapped"
  meaning: ForcefulMotionAction
    constraints
      self.m.x-net  ⟵  Slap-X-Net
```

Figure 6.2. ECG representations of processes related to force application including ForceTransfer, ForceApplication, and ForcefulMotionAction. These schemas represent the meaning of verbs of impact such as "hit", "strike", "kick", "slap", etc.

The meaning block of the hit and slap lexical constructions is typed as ForcefulMotionAction. The meaning block of a construction is quite similar to a semantic schema, and thus it also allows for semantic constraints as well as evoking structure. In the case of the lexical HitPastTense and SlapPastTense constructions, the (inherited) x-net role in ForcefulMotionAction is assigned the appropriate X-Net. Using the general ForcefulMotionAction schema to represent the meaning of "hit" and "slap" provides a useful level of generalization over a broad range of verbs, and, as we will see below, it enables the definition of a simple transitive argument structure construction to cover this semantic class of verbs.

ECG argument structure constructions must also provide guidance about how a scene should be simulated. For example, active and passive provide differing perspectives on the same scene, and such a perspective shift must be communicated to the simulator. For this, the EJ1 grammar uses a general VerbPlusArguments construction and its associated abstraction over scenes (events) called the EventDescriptor schema (both shown in Figure 6.3) to represent perspectivized facts about a scene.

```
schema EventDescriptor
 roles
   eventType: Process
   profiledProcess: Process
   profiledParticipant
   spatialSetting
   temporalSetting
   discourseSegment: DiscourseSegment
```

```
general construction VerbPlusArguments
 subcase of VP
 constructional
   constituents
     v: Verb
 meaning
   evokes EventDescriptor as ed (inherited)
   constraints
     self.m ⟷ ed.eventType (inherited)
     v.m ⟷ ed.profiledProcess
```

Figure 6.3. The EventDescriptor Schema and a general VerbPlusArguments construction that functions as the root of the argument structure hierarchy

The central importance of the EventDescriptor schema extends the central function of predication in grammar as proposed in Croft's *Radical Construction Grammar* (2001). We suggest that an utterance has two primary construals available to it: the scene provided by the argument structure construction and a particular process provided by the verb. Often these two processes are the same, but they are not required to be the same.

In the EventDescriptor schema:

- The eventType role is bound to the Process that represents the scene being described. An argument structure construction supplies the filler of this role.
- The profiledProcess role is bound to the process or subprocess in the scene that is being profiled. The verb supplies the filler of this role.
- The profiledParticipant role is bound to the focal participant in the scene that is being described. This role can be thought of as the semantic correlate of subject and it is bound to different roles in a scene depending on whether the utterance is active or passive voice.
- Roles temporalSetting and locativeSetting are bound to the time and location.
- The discourseSegment role is typed to a simplified DiscourseSegment schema, which has roles for the speechAct of the utterance and the topic of the utterance. In this chapter, the speechAct role will be bound to a simple atomic value such as "Declarative" or "WH-Question". The topic role will be bound to the topic specified by each finite clause. Table 6.1 shows how the roles of the EventDescriptor are co-indexed for some examples.

The profiledParticipant role provides a lot of leverage in the grammar. It allows for a simple semantic distinction between active and passive sentences and makes it straightforward to implement the semantics of control (described below). As a consequence, the subject constituent is just like any other constituent in the grammar, and has no special status apart from the fact that its meaning is bound to the profiledParticipant role. Thus a construction like the imperative without a subject is

Table 6.1. Different ways the EventDescriptor.profiledParticipant, DiscourseSegment.topic, and the semantic roles of CauseEffectAction are co-indexed for four different example sentences

Example	Bindings
He hit the table	profiledParticipant ⟷ topic ⟷ causer
His hand hit the table	profiledParticipant ⟷ topic ⟷ instrument
The table was hit	profiledParticipant ⟷ topic ⟷ affected
Which table did he hit	profiledParticipant ⟷ causer topic ⟷ affected

not problematic in that the profiledParticipant is just bound to the addressee of the utterance.

The notion of topic is distinct from both the subject constituent and the profiledParticipant role. While a simple declarative construction (described below) would co-index the profiledParticipant with the topic role, this is not a requirement. A WH-Question with a fronted NP constituent binds the meaning of the fronted NP to the topic role. Again, the ECG EventDescriptor formalizes and extends several insights from CL. By formalizing theoretical findings from CL, ECG allows for conceptual compositions that model human language use.

Also shown in Figure 6.3 is the phrasal VerbPlusArguments construction. *Phrasal* constructions share many properties with *lexical* constructions. The primary difference between the lexical and phrasal constructions presented in this chapter is the presence of a **constructional** block. A phrasal construction's **constituents** are defined in the constructional block.

The VerbPlusArguments construction is the root of the argument structure hierarchy. Because it is marked with the **general** keyword, it is not used directly for interpretation or production, but instead represents a generalization over all its subtypes. The generalization that the VerbPlusArguments captures is that all argument structure constructions (in this EJ1 grammar) have a verb constituent called v.

The VerbPlusArguments construction has no additional form constraints to add, and thus the form block is omitted. In its meaning block, the VerbPlusArguments construction inherits an evoked EventDescriptor from general construction VP (not shown) as well as the constraint that the meaning of the construction itself (specified by self.m) is bound to the EventDescriptor's eventType role. It then adds the semantic constraint that the meaning of the verb is bound to the EventDescriptor's profiledProcess role. This constraint formalizes the intuition described above that a verb profiles a particular process associated with the scene being described.

Specific subtypes of VerbPlusArguments are shown in Figure 6.4. These argument structure constructions define transitives with a salient causer. The meaning of these transitives is defined using the CauseEffectAction schema. CauseEffectAction is

```
schema CauseEffectAction
 subcase of ComplexProcess
 roles
  process1: ForceApplication
  causer
  affected
 constraints
  protagonist ⟷ causer
  protagonist2 ⟷ affected
  process1.actedUpon ⟷ affected
```

```
construction TransitiveCEA
 subcase of VerbPlusArguments
 constructional
  constituents
   np: NP
 form
  constraints
   v.f before np.f
 meaning: CauseEffectAction
  constraints
   self.m.causer ⟷ ed.profiledParticipant
   self.m.affected ⟷ np.m
   self.m ⟷ v.m
```

```
construction TransitiveCEAProfiledCause
 subcaseof TransitiveCEA
 meaning: CauseEffectAction
  evokes ForcefulMotionAction as fma
  constraints
   ignore : self.m ⟷ v.m
   v.m ⟷ fma
   self.m.process1 ⟷ fma.process1
```

```
construction TransitiveCEAProfiledInstrument
 subcase of TransitiveCEAProfiledCause
 meaning: CauseEffectAction
  constraints
   ignore: self.m.causer ⟷ ed.profiledParticipant
   self.m.process1.instrument ⟷ ed.profiledParticipant
```

Figure 6.4. A selection of argument structure constructions that model transitives like *he cut the steak* (TransitiveCEA), *he hit the table* (TransitiveCEAProfiled-Cause), and *his hand hit the table* (TransitiveProfiledInstrument)

a complex process in which process1 (the cause) is a ForceApplication, and process2 is the effect. It also adds roles for a causer and an affected participant, and uses co-indexations to bind the causer and affected roles to the appropriate roles in process1 and process2.

The TransitiveCEA (Cause Effect Action) construction represents transitive VPs with causal verbs with a force application component such as "cut", "chop", "crush", etc. It defines an additional NP constituent to represent its object, and uses the form block to add a form constraint requiring that the verb's form (specified by v.f) comes before the form of the np (specified by np.f). The meaning of the construction is defined to be the CauseEffectAction schema.

In the semantic constraints block of TransitiveCEA, the construction specifies how its semantic roles relate to its constituents. Its semantic roles are referred to using the slot chain self.m, which in this case refers to a CauseEffectAction schema. Thus self.m.causer refers to its semantic causer role. The first constraint in the meaning block of TransitiveCEA co-indexes the causer with the profiledParticipant, and the second constraint co-indexes the affected participant role with the meaning of the NP constituent. The final constraint unifies the meaning of the construction itself (self.m) with the meaning of the verb. The first consequence of this co-indexation is that it limits the kinds of verbs that can co-occur with the TransitiveCEA construction to those that have meaning poles that can unify with CauseEffectAction. The second consequence is that it imports the meaning of the verb into the meaning pole of the construction via unification.

6.5 Analysis examples

Having introduced some of the schemas and constructions in EJ1, we now show how they are used to support a compositional analysis of various types of sentences. Our emphasis here is on the mechanisms of ECG; the linguistic analysis of these examples and many more can be found in Dodge (2009).

We assume the following points: (1) A sentence not only instantiates lexical constructions but also phrasal and other types of constructions. All of these constructions are meaningful, with meaning represented using schemas. (2) When these constructions unify, their meanings compose in a way that is consistent with the constraints specified in each construction. (3) In ECG, this composed meaning is represented as a semantic specification for a simulation (semspec).

The same construction can be instantiated in many different sentences, and should therefore compose with a variety of other constructions. For each sentence, the instantiated constructions should unify to produce a semspec that is consistent with our intuitions about that sentence's meaning. Similarities and differences in sentence meaning should be reflected in their semspecs. In addition, the ECG lattice of constructions facilitates expressing generalizations across constructions.

We will first present an in-depth analysis of the simple declarative sentence *He hit the table*. Then, we will look at sentences which are similar in many respects, but which present some challenges to linguistic analysis, such as instrument subjects, passives, and questions. For each, we will describe how the EJ1 grammar supports an analysis involving many of the same or similar constructions composed in different ways. Crucially, differences in meaning are captured as a few key differences in the semspecs that result from these different compositions.

He hit the table instantiates several different constructions, whose meanings are unified to produce the semspec shown in Figure 6.6. To understand which elements of the semspec each of these constructions provide, how they unify, and what

```
general construction S-WithSubj
 subcase of S
 constructional
   constituents
     subj: NP
     fin: Finite
 meaning: EventDescriptor
   constraints
     subj.m ⟷ self.m.profiledParticipant
```

```
construction Declarative
 subcase of S-WithSubj
 constructional
   constituents
     fin: FiniteVP
 form
   constraints
     subj.f before fin.f
 meaning: EventDescriptor
   constraints
     self.m ⟷ fin.ed
     self.m.discourseSegment.speechAct ⟵ "Declarative"
     self.m.discourseSegment.topic ⟷ self.m.profiledParticipant
```

Figure 6.5. Speech act constructions that set the profiledParticipant role and the topic roles of the EventDescriptor and DiscourseSegment schemas, respectively. The Declarative construction covers basic declarative sentences

type of sentence meaning this semspec represents, let's look more closely at the instantiated constructions. The key ones described here are: Declarative (shown in Figure 6.5), TransitiveCEAProfiledCause (an argument structure construction, shown in Figure 6.4); and HitPastTense (a verb construction shown in Figure 6.2). In addition, there are nominal constructions for "he" and "the table".

Declarative identifies its meaning with an EventDescriptor schema (Figure 6.3), indicating that this type of construction is used to describe some kind of event. Declarative inherits a subj constituent and the constraint that this constituent's meaning is bound to the EventDescriptor's profiledParticipant role. In this way, Declarative indicates that the event should be simulated from the perspective of the entity referred to by the subj constituent. Constituent subj, in the current example, unifies with HE, providing information that the entity that fills the profiledParticipant role is, in this case, MALEANIMATE.

Declarative does not, however, specify what type of event is being described, nor does it specify which event-related semantic role the profiledParticipant

$$
\left[\begin{array}{ll}
\textbf{EventDescriptor} & \\
\text{discourseSegment:} & \left[\begin{array}{ll}\textbf{DiscourseSegment} & \\ \text{speechAct:} & \text{"Declarative"} \\ \text{topic:} & \boxed{2}\end{array}\right] \\
\text{profiledParticipant:} & \boxed{2} \\
\text{eventType:} & \left[\begin{array}{ll}\textbf{CauseEffectAction} & \\ \text{causer:} & \boxed{2}\ \text{MALEANIMATE} \\ \text{process1:} & \boxed{4}\left[\begin{array}{ll}\textbf{ForceApplication} & \\ \text{instrument:} & \boxed{3} \\ \text{actedUpon:} & \boxed{1}\ \text{TABLE} \\ \text{protagonist:} & \boxed{2}\end{array}\right] \\ \text{x-net:} & \boxed{5} \\ \text{process2:} & \left[\begin{array}{ll}\textbf{Process} & \\ \text{protagonist:} & \boxed{1}\end{array}\right] \\ \text{protagonist:} & \boxed{2} \\ \text{affected:} & \boxed{1} \\ \text{protagonist2:} & \boxed{1}\end{array}\right] \\
\text{profiledProcess:} & \left[\begin{array}{ll}\textbf{ForcefulMotionAction} & \\ \text{process1:} & \boxed{4} \\ \text{x-net:} & \boxed{0}\ \text{Hit-X-Net} \\ \text{process2:} & \left[\begin{array}{ll}\textbf{EffectorTranslationalMotion} & \\ \text{process1:} & \left[\begin{array}{ll}\textbf{MotorControl} & \\ \text{effector:} & \boxed{3}\end{array}\right] \\ \text{target:} & \boxed{1} \\ \text{protagonist:} & \boxed{2}\end{array}\right] \\ \text{protagonist:} & \boxed{2} \\ \text{protagonist2:} & \boxed{2}\end{array}\right]
\end{array}\right]
$$

Figure 6.6. The semspec for the sentence *He hit the table*

is associated with; this information is instead supplied by whichever argument structure construction Declarative unifies with, in this case TransitiveCEA ProfiledCause.

Declarative also specifies discourse-related information. This is represented here in simplified fashion by specifying that the discourse segment's speechAct role has the atomic value "Declarative". Additionally, the topic of the discourse segment is bound to the profiledParticipant role. Thus, the profiledParticipant, the topic of the discourse segment, and the meaning of the subj constituent are all bound to one another, as shown by their sharing boxed number 2 in Figure 6.6.

Declarative has a second constituent, fin (a type of FiniteVP, not shown) which, in this example, unifies with TransitiveCEAProfiledCause. This argument structure construction unifies with verbs such as "hit", "slap", and "kick". Declarative also specifies that the EventDescriptor evoked by this argument structure construction is to be identified with that of Declarative, indicating that both constructions are describing the same event.

As with other argument structure constructions, TransitiveCEAProfiled Cause provides information about the type of scene that is being described. This is specified through the (inherited) constraint that its meaning is bound to the eventType role of the EventDescriptor. TransitiveCEAProfiledCause is a member of a family of transitive argument structure constructions, all of which identify their meaning with a CauseEffectAction schema (Figure 6.4). This schema defines a complex process in which one process, a ForceApplication, has a causal relation to another process. Two key participant roles are defined by this schema: an animate causer who performs the action, and an affected entity, which is acted upon and (potentially) affected in some way by this action. Thus, this family of constructions reflects the causal semantics prototypically associated with transitivity, and includes semantic roles similar to those of Agent and Patient.[2] In ECG, both the representation and underlying conceptualization of these roles is semantically complex, and they are defined relative to embodied schemas/gestalts, rather than just being names whose meaning is not explicitly specified.

In EJ1 grammar, argument structure constructions not only specify what type of event is being described but also specify which event participant is "profiled", i.e., from whose perspective the event should be simulated. Transitive CEAProfiledCause is a subcase of the TransitiveCEA, and inherits its constraint that the profiledParticipant is the causer. Therefore, in the semspec, causer is co-indexed with profiledParticipant. As described later, other argument structure constructions may specify that the profiledParticipant is bound to a different event role.

As with other argument structure constructions, TransitiveCEAProfiled Cause inherits a verb constituent, the meaning of which serves to elaborate a process or subprocess related to the event as a whole. In the semspec, the meaning of the verb is bound to the profiledProcess of the EventDescriptor. Its parent construction, TransitiveCEA, defines a central case situation in which the verb constituent meaning is identified with the same schema as the argument structure construction, indicating a very close correspondence in meaning between the two constructions. TransitiveCEAProfiledCause represents an extension to the central case, a situation in which the verb constituent provides information about the causal process of CauseEffectAction (process1), but does not elaborate the effect (process2). This is specified within the construction by: (1) constraining verb constituent meaning to be ForcefulMotionAction and specifying that the inherited constraint that verb meaning is the same as the argument structure construction meaning should be ignored, and (2) binding the causal subprocess, ForceApplication, to the verb constituent meaning. Because constructional meaning is specified using conceptual primitives, this argument structure construction is not lexically-specific, covering all verbs which identify their meaning with

[2] Note, though, that not all argument structure constructions of transitive form necessarily share this same meaning.

ForcefulMotionAction (such as "punch", "pat", and "tap"). The unification of the verb and argument structure constructions results in the co-indexation of many different roles.

In Figure 6.6, the causer (boxed 2) is also protagonist of CauseEffect Action, ForceApplication, ForcefulMotionAction, and EffectorTranslationalMotion. The affected (boxed 1) is:

- CauseEffectAction.protagonist2
- CauseEffectAction.process2.protagonist
- ForceAppplication.actedUpon
- EffectorTranslationalMotion.target.

TransitiveCEAProfiledCause's verb constituent is unified with the HitPastTense verb construction, whose meaning (ForcefulMotionAction) meets the constraints specified by the argument structure construction. HitPastTense specifies a particular type of X-net (a hitting routine). The NP constituent of TransitiveCEAProfiledCause is bound to the affected role. It provides information about the affected role of the CauseEffectAction. In this particular example, this constituent is unified with an NP construction whose N constituent is TABLE.

The semspec for this sentence thus supports the simulation of an event in which a male causal actor performs a particular kind of forceful action (hitting) on a table, and this table is (at least potentially) affected in some way. Neither the argument structure construction nor the verb specify what type of effect this is; the simulated effect will depend on the particular fillers of the causer and affected roles. Compare: The baby/weightlifter hit the table/wine glass/policeman. The simulation of effect will also depend on the specific ForcefulMotionAction described by the verb. For instance, substitute *patted*, which specifies a low amount of force, for *hit* in the examples above.

We have now shown how a particular sentence can be analyzed as instantiating several constructions, whose meanings unify to produce a semspec for that sentence. Next, we will show how sentences similar in form and meaning to *He hit the table* can be analyzed as compositions of many—but not all—of these same constructions.

First, consider the sentence *The hammer hit the table*. Under an instrumental reading[3] of this sentence, the event being described is one in which a person used a hammer to hit the table—roughly, someone hit the table with a hammer. The meaning of this sentence is very similar to *He hit the table*, but with an important difference in participant profiling: *The hammer hit the table* foregrounds the

[3] Note alternative reading = non-agentive impact, e.g., *The hammer fell and hit the table*. This reading would also focus on the hammer's motion and contact, but the actor and actions would no longer be part of the conceptual picture. This reading would be analyzed as a different type of event, one which does not include an agentive causer.

instrument and its motion and contact with the table, and backgrounds the actor who is wielding this instrument.

In ECG, the particular "event perspective" or "participant profiling" of an utterance is specified in the semspec through a binding between profiled Participant and a particular event participant role. In the case of *He hit the table*, profiledParticipant is bound to causer. But, for *The hammer hit the table*, this role is bound to an instrument role instead (see Table 6.1).

This distinction is specified in EJ1 by using a different, but closely related, argument structure construction. *He hit the table*, instantiates TransitiveCEA ProfiledCause, which specifies that the profiledParticipant is bound to the causer role. *The hammer hit the table* is analyzed as instantiating a subcase of the argument structure construction: TransitiveCEAProfiledInstrument (see Figure 6.4). This subcase **ignore**s the parent's specification that profiled Participant is bound to causer, and instead specifies that it is bound to the instrument role of the ForceApplication action (i.e., the instrument the actor used to apply force to the actedUpon). But, these argument structure constructions are the same in all other respects.

The other constructions instantiated in this example are the same, with the exception of the particular NP that is bound to Declarative's subj constituent. Recall that Declarative specifies that the meaning of its subj constituent is bound to profiledParticipant. Unification therefore results in co-indexation of the meaning of this NP construction with the semantic role bound to profiledParticipant, which in this case is the Instrument of ForceApplication. The resultant semspec therefore would differ from Figure 6.6 in that profiledParticipant is co-indexed with an Instrument role rather than the causer role. Such a semspec indicates that the event should be simulated from the perspective of the instrument, not the causer. Furthermore the semspec would indicate that instrument is of type HAMMER, information not present in previous example. At the same time, we lose information that causer is MALEANIMATE, since this role is no longer explicitly expressed. However, the role itself is still part of the event description, indicating that the causer is still conceptually present.

6.6 Extensions

These examples just discussed differ as to which argument structure constructions they instantiate. But, in both cases, the argument structure construction composes with an instance of Declarative, indicating that in both cases the kind of speech act is the same, and the subject is also the discourse topic. Of course to apply to a broader range of linguistic data, these argument structure constructions must also compose with constructions that provide different discourse-related specifications.

Questions—Which table did he hit?

To illustrate how this can be done, consider the sentence *Which table did he hit?*, which differs from *He hit the table* in terms of the type of speechAct. Both sentences have the same subject, but differ as to discourse topic. However, the meanings of both sentences are similar with respect to the type of event being described and the perspective from which it is described.

Which table did he hit? instantiates the same verb and argument structure constructions as *He hit the table*, as well as similar nominal constructions. When the instantiated constructions unify, the resultant semspec is very similar to that of *He hit the table*. Both specify the same eventType and profiledProcess, and in both profiledParticipants is co-indexed with causer. The key differences relate to the DiscourseSegment roles. Firstly, speechAct types are different ["WH-question" vs. "Declarative".] And secondly, the topic is co-indexed with different semantic roles in each case: for declarative, the topic is co-indexed with causer (and meaning of subj constituent) whereas in the WH-question the topic is co-indexed with affected (and the meaning of the extraposed constituent). The best-fit analyzer of section 6.7 hypothesizes that the extraposed phrase can fill the affected role of the TransitiveCEAProfiledCause construction because there is a good form and meaning fit.

These analyses are possible, in part, because the EventDescriptor schema is defined such that the notion of topic is separate from both the subject constituent and the profiledParticipant role. In Declarative, these conceptual elements are all bound to one another, indicating a particular type of conceptual whole. But, because this particular combination is not represented by an atomic role it is also possible to write constructions in which these are not all bound, such as the question construction instantiated in *Which table did he hit*? EJ1 argument structure constructions are defined such that they can compose not only with declarative but also question (and other types of speech act) constructions.

Passives—The table was hit (by him).

The passive is typically analyzed in relation to active, but the exact relation remains a topic of continuing linguistic research. In the analysis that we sketch here, actives and passives are treated as different families of constructions which are related through common semantics. The basic idea is to have passive constructions use the exact same schemas as their active counterparts, while inheriting their form constraints through the passive hierarchy.

Using EJ1, *The table was hit (by him)* is analyzed as instantiating a passive argument structure construction whose meaning is CauseEffectAction, and which specifies that the meaning of its verb constituent is one of ForcefulMotion Action. In these respects, this construction is the same as the Transitive CEAProfiledCause argument structure construction that was instantiated in *He hit the table*. However, the passive argument structure construction specifies that the profiledParticipant role is bound to the affected role of the CauseEffectAction, not the causer role. Thus, both the passive and the active argument structure constructions specify the same type of event, and both have a verb constituent that elaborates the causal action within this event. But, they differ on which simulation perspective they specify, with active specifying that of the causer, and passive that of the affected.

Constructionally, the passive differs from active in terms of its constituents. Unlike TransitiveCEAProfiledCause, the passive argument structure construction does not have an NP constituent, but does have an optional prepositional phrase, whose meaning is bound to the causer role. In addition, passive has different verb constituent constraints, including the fact that the verb form is that of past participle.

The table was hit therefore instantiates a different (though semantically similar) argument structure construction from *He hit the table*. Most of the other instantiated constructions are the same for both examples, including Declarative, a HitPastTense verb construction, and an NP construction for "the table". When these constructions are unified, the semspecs are also very similar, with the key difference that in *The table was hit* the profiledParticipant and topic is the affected participant rather than the causer (see Table 6.1).

Control—He wanted to hit the table.

The strategy for handling control relations in ECG also relies on leveraging the power of the profiledParticipant role. The basic idea can be illustrated by a description of the strategy for analyzing the sentence *He wanted to hit the table*, as follows.

First, define a set of control verbs, such as *want*, whose meaning can be defined as involving an additional "event" role. The meaning of *want*, for example, would be represented as a "wanting" process in which a wanter desires that some type of event take place.

Next, define an argument structure construction whose verb constituent is a control verb, and which adds another constituent that is itself an argument structure construction. Thus, there will be two EventDescriptors in the semspec, one associated with the control argument structure construction itself, and the other associated with the constituent argument structure construction. For the current example, this second argument structure construction would be TransitiveCEAProfiledCause, the same argument structure construction instantiated in *He hit the table*. In the resulting semspec, the control EventDescriptor describes the event of wanting something, and the second EventDescriptor describes the thing that is wanted, which in this case is a hitting event. In addition, the profiledParticipant role of the main EventDescriptor is bound to the control verb's protagonist role, which in the current example is the "wanter" role.

For the subject control argument structure construction instantiated in this example, the profiledParticipant is also bound to the profiledParticipant of the constituent argument structure construction. Therefore, in the semspec for *He wanted to hit the table*, the profiledParticipants of each event descriptor are co-indexed, and are co-indexed with the wanter of the wanting event, and the causer of the "hitting the table" event. And, because *He wanted to hit the table* also instantiates Declarative, profiledParticipant is also be bound to the meaning of the subj constituent "He".

6.7 Analyzer/workbench

The analyses described in this chapter are produced by a system called the constructional analyzer (Bryant 2008). Constructional analysis is the process of interpreting an utterance in context using constructions, and the analyzer maps an utterance onto an instantiated set of ECG constructions and semantic schemas. The design of the system is informed by the fields of construction grammar/functional linguistics, natural language processing and psycholinguistics, and the constructional analyzer is a cognitive model of language interpretation within the tradition of Unified Cognitive Science and NTL.

The power of the analyzer comes from combining constructions with best-fit processing. Best-fit is a term we use to describe any decision-making process that combines information from multiple domains in a quantitative way. Thus best-fit constructional analysis is a process in which decisions about how to interpret an utterance are conditioned on syntactic, semantic, and contextual information. Because constructions provide explicit constraints between form, meaning and context, they are well-suited to a best-fit approach (Narayanan and Jurafsky 2001).

The best-fit metric computes the conditional likelihood of an interpretation given the grammar and the utterance and is implemented as a factored probabilistic model over syntax and semantics. The syntactic factor incorporates construction-specific preferences about constituent expression/omission and the kinds of constructional fillers preferred by each constituent. The semantic factor scores a semspec in terms of the fit between roles and fillers.

The constructional analyzer uses a psychologically plausible sentence processing algorithm to incrementally interpret an input utterance. Each partial (incremental) interpretation is a subset of the instantiated constructions and schemas that go into the final, intended interpretation. Intuitively, this means that there is a set of competing partial interpretations that are each trying to explain the prefix of the input that has been seen so far. The best-fit metric is used to focus the analyzer's attention on more likely partial interpretations.

The analyzer produces rich linguistic analyses for a range of interesting constructions including embodied semspecs for the various motion and force-application constructions designed by Dodge (2009). A variety of syntactically interesting constructions are also easy to implement within the analyzer including constructions for passives, simple WH-questions, raising and a radial category description of the ditransitive argument structure construction.

Although the English construction grammar is currently the most linguistically well-motivated grammar processed by the analyzer, the analyzer is not tied to English. It analyzes Mandarin child-directed utterances as well, using a Mandarin grammar. Productive omission is incorporated into the system and scored by the best-fit metric (Mok and Bryant 2006). Omitted arguments are resolved to a candidate set by a simple context model.

The analyzer also predicts differences in incremental reading time for reduced relative data. The factored syntactic and semantic model plays an important role in making the reading time predictions. The syntactic factor implements the syntactic bias for main verb interpretations over reduced relative interpretations, and the semantic model implements the good agent/good patient biases that differentiate the two experimental conditions (Bryant 2008). Crucially, the system is a model of deep semantic interpretation first, and it predicts reading time data as a by-product of its design. Luca Galiardi's ECG workbench provides extensive support for grammar building and testing and is further described on the ECG wiki.

6.8 Conclusions

The main purpose of this chapter is to introduce both the technical aspects of ECG and the scientific basis for the underlying NTL. By presenting detailed examples, we hope to convey how ECG and the related tools can be used for deeper linguistic analysis than is otherwise available. Even from the fragments presented here, it is clear that ECG grammars employ a large number of constructions, contrary to traditional minimalist criteria for language description. This is partly a question of style, as one could define an equivalent formalism that had fewer, but more complicated (parameterized) constructions. More basically, we believe that constructional compositionality crucially depends on a deep semantics that captures the rich structure of human conceptual systems. The semantic poles of ECG constructions are based on our best understanding of the conceptual and communicative primitives. In addition, we suggest that the critical resource in language processing is time, not space. By having explicit, compositional, constructions for language variations we simplify grammar writing and analysis for the analyzer program and, we believe, this is true for human language processors as well.

The NTL project and the ECG grammar formalism are both undergoing continuous development and this chapter presents only a snapshot of one thread. Below, we outline some of the current areas of research. John Bryant's dissertation (Bryant 2008) contains more information on all of this work.

Constructions should be able to capture form–meaning relations at all levels of language use. We have extended ECG and the analyzer to handle complex morphology, including Semitic, by constructing a parallel morphological analyzer that coordinates with the system described here. The idea of a common deep semantics linking various forms of time-locked input is, in principle, extendable to speech, intonation, and gesture.

There are now pilot implementations of two additional ECG primitives, *situations* and *maps*, which are needed to handle key CL mechanisms such as metaphors and mental spaces. As part of this extension we are also incorporating techniques for modeling language communities and social communication.

This chapter, and the bulk of NTL work, has focused on language recognition; modeling *production* brings in a wide range of additional issues of audience modeling, etc. Interestingly, the best-fit analyzer already does analysis-by-synthesis and would not require major redesign to generate the best surface form, given metrics on the desiderata.

One of the most ambitious current projects involving ECG is an attempt to model in detail how children acquire their early grammatical constructions. Because of its explicit linking of embodied conceptual structure to linguistic form, ECG seems to provide a uniquely appropriate foundation for such studies.

All inductive learning is statistical, but the NTL work differs from purely statistical studies in postulating some conceptual and grammatical primitives as the hypothesis space for learning. The conceptual primitives include all of the embodied concepts (including emotional, social, etc.) that the child brings to language learning. The grammatical prior consists of three basic assumptions:

(a) The child knows many meaning (conceptual) relations
(b) The child can recognize relations in language form (e.g., word order)
(c) Grammatical constructions pair form relations with meaning relations.

Since the primitive relations in both form and meaning are bounded, the learning problem for the child (and our computer models) is not intractable. Ongoing work by Nancy Chang and Eva Mok (Chang and Mok 2006) demonstrates that

ECG-based programs can learn complex grammatical constructions from labeled input, even for languages like Mandarin that allow massive omission.

We have said relatively little in this chapter about the neural realization of our Neural Theory of Language. A great deal of ECG-based linguistic analysis can be done without explicit neural considerations, just as much biology can be done without chemistry. But the neural perspective is crucial for many aspects including developing testable models and conceptual primitives. Our idea of how all the levels integrate is presented in Feldman (2006), as part of Unified Cognitive Science.

As we suggested at the beginning, this growing Unified Cognitive Science presents opportunities of new possibilities for deep semantic grammars for theoretical, scientific, and practical uses. When we add powerful tools, such as those described in this chapter, the future of linguistic analysis looks very promising.

This page intentionally left blank

CHAPTER 7

SIGN-BASED CONSTRUCTION GRAMMAR

LAURA A. MICHAELIS

7.1 Introduction

To practice Construction Grammar is to accept a proposition that is anathema to most linguists, whether they be "formalists" or "functionalists": many, if not most, of the grammatical facts that people appear to know cannot be resolved into general principles—whether these concern semantics, information processing, or conversational practice—but must instead be stipulated. This stipulation takes the form of a grammatical construction. Grammatical constructions are recipes for word combinations that speakers use to achieve specific communicative goals, e.g., issuing an order, requesting information, attributing a property to an entity. Constructions determine the linear order of the words—as the English verb–phrase construction requires the direct object to follow the verb—and the forms of the words—as the comparable Latin construction requires its direct object to have an accusative case ending. A construction cannot be pronounced, but, like a word, it is a conventionalized pairing of form and meaning (Fillmore et al. in prep.; Goldberg 1995; Michaelis and Lambrecht 1996; Croft 2001). In viewing syntactic patterns as meaningful, Construction Grammar represents a significant departure from phrase-structure-based theories of grammar. In standard generative theory, rules of grammar create word combinations that express composite concepts like predicates and propositions, but these rules do not add any meaning to that contributed by the individual words. Thus, on the standard view of syntax, phrases have meaning but the rules that create phrases do not. On the constructionist view, phrasal patterns not only have meanings but also can shift the meanings of the words that they combine. In fact, such shifts constitute one of our major lines of evidence for the existence of constructions (Michaelis 2004). As a simple example, consider the rare but attested denominal verb *sister*, as in *We sistered the joints*. In this context the word has a causative interpretation ("cause two things to be sisters"). Fully understanding the meaning of this word in this sentence requires knowledge of the noun *sister*, an image-based metaphorical mapping and perhaps some background in carpentry, but the interpretive affordance that this word represents exists only in the context of the transitive VP pattern.

To establish the need for a particular construction (e.g., the transitive VP), one need only show that independently motivated principles fail to predict all of the facts about the use, internal composition, combinatoric potential, or meaning of the pattern under study. Thus it could be said that constructionists enjoy a lower burden of proof than other syntacticians: like a defense attorney, a construction grammarian need only cast reasonable doubt on the opponent's theory of the case, however coherent and compelling it may be, to win. But what in fact does the constructionist win? Construction-based syntax, at least as it was practiced in the 1960s and 1970s, is widely regarded as a failed experiment:

> Although syntactic work within the transformationalist tradition frequently uses the term descriptively, "(grammatical) construction" has been a theoretical taboo at least since the 1980s. Briefly, Chomsky argued that transformations like "passive" and "raising", common in earlier versions of transformational grammar, could be eliminated in favor of general conditions on structures that would allow a single operation—Move NP—to do the work of a family of such transformations. This has guided the subsequent evolution of transformational analysis where one now finds discussion of even more general operations, such as "Move α" or "Move". This evolution has tended to move away from construction-specific proposals toward a discussion focused almost exclusively on general principles from which the idiosyncrasies of individual constructions are supposed to be derived.
>
> (Ginzburg and Sag 2004: 4)

Certainly, construction-based transformational grammars lacked a satisfying way to express cross-constructional generalizations: for example, each unbounded movement transformation specified the same movement operation operating over the same unbounded context as every other such transformation. But transformational approaches to grammar lacked this capacity precisely because they did not consider grammatical patterns, like relative clauses, information questions, and topicalization, to be units of grammar. Transformational grammar was designed to represent one type of relationship—that between tree structures—and tree structures are not in grammar. Instead, they are created online through recursive application of phrase-structure rules. The recognition that many transformations, including "dative movement" and passive, are "lexically triggered" (restricted to certain classes of lexical items) caused proponents to replace a number of transformations with lexical rules, which place lexical entries into correspondence. But neither lexical rules nor transformations could do their work without a considerable number of provisos, necessary to account for both lexical exceptions and pieces of structure that transformations must somehow introduce. As an example of a lexical exception, consider those verbs like *ask*, which, while welcoming the ditransitive or double-object frame as in (1), do not occur in the putative input frame, the oblique-goal pattern, as in (2):

(1) They asked me a question.

(2) *They asked a question to me.

As an example of a structure-adding transformation, consider the passive-voice transformation, whose input and output structures are exemplified in (3–4), respectively:

(3) The committee discussed the proposal.

(4) The proposal was discussed by the committee.

Here is passive, as described by Ginzburg and Sag (2000):

> [As] noted by McCawley (1988) in his review of Chomsky (1986*a*), Chomsky's discussion of the passive construction did not touch on crucial issues like the relevant verb morphology, the choice of the preposition *by*, and the role of the verb *be*. As McCawley pointed out, these properties of the construction followed from nothing under Chomsky's proposals. Rather, they would have to be stated in a fashion that would render Chomsky's proposal comparably stipulative to the alternative it sought to replace. (Ginzburg and Sag 2000: 4)

If stipulation is required anyway, the reasoning goes, there is no reason to retain transformations and other mapping procedures, and a good reason to eliminate them: since procedures are not grammar objects, they have no ontology. Constructions, by contrast, are objects (or, more accurately, descriptions of objects); as such, they are subject to typing and taxonomic organization. The idea that syntactic rules can be made amenable to taxonomic organization—an idea that links Construction Grammar to an allied theory, Head-Driven Phrase-Structure Grammar (Sag et al. 2003)—has been central to Construction Grammar argumentation from its earliest incarnations (Fillmore et al. 1988 and Lakoff 1987). The taxonomies (called *inheritance hierarchies*) are offered as tools for describing shared semantic, pragmatic, and grammatical properties of syntactic patterns, in much the same way that category members are said to be linked by family-resemblance relations (Lakoff 1987). If there is a theme running through all construction-based syntactic research it is this: we do not sacrifice linguistic generalizations by stipulating idiosyncratic properties of constructions because detailed constructions are instances of more abstract constructions.

One can in fact view construction-based theories of syntax as upholding standards of grammar coverage that the original proponents of generative grammar have abandoned, as they sought to reduce the theory's dependence on linguistic facts:

> A look at the earliest work from the mid-1950s will show that many phenomena that fell within the rich descriptive apparatus then postulated, often with accounts of no little interest and insight, lack any serious analysis within the much narrower theories motivated by the search for explanatory adequacy, and remain among the huge mass of constructions for which no principled explanation exists—again, not an unusual concomitant of progress.
> (Chomsky 1995: 435)

It seems safe to say that most proponents of construction-based syntax would not consider the loss of insightful grammatical descriptions a mark of progress. Further, it is questionable whether *narrower* properly describes the relationship between Chomsky's program and the formalized version of Construction Grammar to be described in this chapter: Sign-Based Construction Grammar (SBCG). It seems reasonable to assert that a formal theory like SBCG is *ipso facto* "narrower" than an informal one, like the Minimalist Program, if only because formalism imposes a limit on potential predictions. The SBCG formalism will be the focus of the following section. In subsequent sections, I will discuss four rationales that constructionists have offered for a construction-based view of syntax. These are:

- Constructions license arguments and syntactic sisterhood relations (section 7.3)
- There is a continuum of idiomaticity (section 7.4)
- Core and periphery are interleaved during production (section 7.5)
- Constructions have formal and interpretive conditions that cannot be captured by mapping procedures (section 7.6).

In the concluding section, section 7.7, I will discuss the role of construction-based syntax in the search for syntactic universals.

7.2 The history and formal architecture of SBCG

The origins of Construction Grammar Common can be traced to a series of case studies published by Berkeley linguists in the late 1980s. These papers target idiomatic grammatical patterns that, while falling outside the descriptive mechanisms of phrase-structure-based grammar, are nonetheless highly productive.

Among these papers are Lakoff's (1987) study of *there* constructions, Fillmore et al.'s (1988) study of the conjunction *let alone*, and Lambrecht's (1987) study of presentational cleft sentences in spoken French. Each promotes a vision of grammar as a structured inventory of form–meaning pairings and, while providing few formal details, advocates a single-format representation for patterns at all points on the gradient from frozen idiom to fully productive rule. One extension of this tradition is found in Goldberg's seminal work on argument-structure constructions (Goldberg 1995; 2002; 2006), Michaelis and Lambrecht's (1996) analysis of exclamatory constructions, and Michaelis and Ruppenhofer's (2001) analysis of German *be*-prefixation. These works, based in part on Langacker's Cognitive Grammar (1987), focus on patterns of semantic extension in constructional meaning, and the semantic shifts that occur when constructions combine with words. This focus on semantic networks is also present in Croft's (2001) Radical Construction Grammar, which uses event-structure representations as the basis for syntactic typology. Croft, like Van Valin and LaPolla (1997), treats grammatical functions and syntactic categories as construction-specific rather than universal roles.

While the foregoing works focus on the structure of the grammar, other work in the Construction Grammar tradition has focused on concerns closer to the hearts of generative syntacticians: the licensing of word strings by rules of syntactic and semantic composition. This research stream is represented by Fillmore and Kay (1995) and Kay and Fillmore (1999). These works, inspired by Generalized Phrase Structure Grammar (Gazdar et al. 1985), outline a unification-based implementation of Construction Grammar in which the grammar is an inventory of syntactic trees with feature structures (rather than syntactic-category labels) at their nodes. These trees are represented as nested (box-within-box) feature structures, the limiting case of which is a single-node feature structure. Feature structures of the latter type are used to describe lexeme classes (e.g., the ditransitive verb construction). Constructions and lexical items are combined by means of unification, which allows the combination of nonconflicting feature structures. Computationally implemented versions of this formalism designed to articulate with sensory-motor schemas include Embodied Construction Grammar (Feldman et al. this volume) and Fluid Construction Grammar (Steels and De Beule 2006).

Despite strong interest in construction-based grammar within computational and cognitive linguistics, Construction Grammar has had little effect on the way that syntacticians do business. This must be attributed, at least in part, to the fact that Construction Grammar does not yet provide a fully elaborated or axiomatized system of sentence representation. To remedy this situation, some of the original proponents of Construction Grammar have begun to collaborate on a formalized version of the theory, SBCG (Fillmore et al. in prep., Sag 2007; 2008). This is the variety of Construction Grammar that I will focus on in this chapter. In SBCG, a construction is a description of a *construct*, which might intuitively be described as a "local tree". The nodes of the trees in such descriptions are not category labels, as in traditional phrase-structure grammar, but *signs*. The notion of sign employed here is close to that of Saussure (1916): a conventionalized pairing of form and meaning. But in SBCG signs have a specific formal realization. One can think of a sign as a node in a syntactic tree, to which certain syntactic and semantic properties accrue. However, signs are more accurately described as feature structures that specify values for the attributes listed in (5–8):

(5) SYN is used to distinguish signs from one another. Its values are the features CAT and VAL(ENCE). The value of CAT is a syntactic category. The VAL feature represents the objects with which a given sign can combine. The VAL value of pronouns, proper nouns, and most common nouns is an empty list. The VAL value of a verb is its combinatoric potential (e.g., the VAL value of a transitive verb is <NP, NP>).

(6) SEM describes the meaning of a sign; its values are the features INDEX and FRAMES. INDEX is the extension of a sign. The FRAMES feature is used to enumerate the predications that together specify the meaning of a sign. In addition to frames representing classic frame-semantic content (e.g., the semantic roles licensed by verbs and other predicators), SBCG uses frames to represent quantifier meanings; these are referred to as quantifier frames. For example, the meaning of the indefinite article *a* in English is represented by means of an existential-quantifier frame.

(7) FORM is used to specify the morphological properties of a given sign; the value of FORM is a (possibly empty) list of morphological entities.

(8) CONTEXT is used to specify features of context that are relevant to the interpretation and use of a given sign.

The subtypes of sign are *word*, *lexeme*, and *phrase*. According to a principle that Sag (2007) refers to as the *sign principle*, signs are licensed in two ways: by a lexical entry or by a construction. Accordingly, the grammar is viewed as consisting of a lexicon—a finite set of lexical descriptions (descriptions of feature structures whose type is either *lexeme* or *word*) and a set of constructions. In (9) we see an example of a lexeme sign:

(9)
$$\begin{bmatrix} \textit{lexeme} & \\ \text{FORM} & \langle \textit{drink} \rangle \\ \text{SYN}|\ \text{VAL} & \left\langle \text{NP}\begin{bmatrix}\text{overt}\\ \text{INST i}\end{bmatrix}, \text{NP}\begin{bmatrix}\text{(ini)}\\ \text{INST x}\end{bmatrix}\right\rangle \\ \text{SEM}|\ \text{FRAMES} & \left\langle \begin{bmatrix}\textit{drink-fr} & \\ \text{DRINKER} & \text{i} \\ \text{DRAFT} & \text{x}\end{bmatrix}, \begin{bmatrix}\textit{animate-fr} & \\ \text{INST} & \text{i}\end{bmatrix}, \begin{bmatrix}\textit{liquid-fr} & \\ \text{INST} & \text{x}\end{bmatrix}\right\rangle \end{bmatrix}$$

The lexeme represented in (9) is *drink*. The semantic properties of this lexeme are represented by a series of frames (e.g., the frame abbreviated as *drink-fr*).

Frames are used to capture the requirement that the drinker be animate and that the consumed item be a liquid. The combinatoric properties of this lexeme are represented in its valence set, which includes two noun phrases—the first of which is coindexed with the "drinker" participant in the drink semantic frame and the second of which is coindexed with the "draft" participant in the drink frame. In addition, each valence member (or valent) is tagged with a feature that represents its instantiation properties: the first valent (the subject NP) is obligatorily instantiated, while the second is optionally null instantiated. As indicated, the second valent, when null instantiated, has an indefinite or, equivalently, existential interpretation. For example, (10) means something like "She drank *some liquid substance* from a plastic mug" (Fillmore 1986):

(10) She drank from a plastic mug.

Words and lexemes are signs all by themselves, while constructions describe sign combinations, which are called *constructs*, as mentioned in section 7.2. It is important to realize, however, that constructions are not trees, or even descriptions of trees, in the sense of traditional phrase-structure grammar. A construction describes only the mother sign of a construct. This mother sign has no daughters but a *daughters* feature: a list-valued attribute. As an illustration of a construction, consider the subject–predicate construction, as described by Sag (2007):

(11)

$$subjpred\text{-}cxt \Rightarrow \left[\begin{array}{l} \textit{phrase} \\ \text{MTR [SYN [VAL } < >]] \\ \text{DTRS } < \text{X, H} > \\ \text{HD-DTR H} \left[\begin{array}{l} \text{SYN [CAT [VF } \textit{fin}] \\ \text{VAL } < \text{X} > \end{array}\right] \end{array}\right]$$

The subject–predicate construction describes the mother sign of a specific type of phrase, a basic clause. Like all constructions, (11) is an implicational statement. This implicational statement says that if a feature structure is the mother sign of a subject–predicate construct, it will contain a mother (MTR) feature with an empty valence list, a daughters (DTRS) feature with two items on its valence list, and a head daughter (H) that is a finite verb and has one item on its valence list (X). X represents the subject of the clause. Like its close congener Head-Driven Phrase-Structure Grammar (Sag et al. 2003), SBCG models the combinatoric properties of words and their phrasal expansions by means of *valence cancellation*. Predicators like verbs and prepositions have valence sets, a list-valued feature that represents the arguments (participant roles) that the predicator requires. As a predicator is combined with the argument(s) that it seeks, that argument is "crossed off" the predicator's valence list. Thus, the mother sign of a subject–predicate construct has an empty valence list: by definition, such a construct contains a daughter (X) that completes the argument requirements of its head daughter, the predicate.

While traditional generative syntax sees syntax, semantics, and lexicon as independent modules, and characterizes the lexicon as a bag of idiosyncratic particulars, SBCG sees the lexicon as having a taxonomic structure, which is referred to as an *inheritance hierarchy* or *type hierarchy*. The items that are organized by such a hierarchy are signs, or, equivalently, feature structures. Signs have a taxonomic organization because each sign belongs to several different grammatical types at once. For example, the verb *discusses* belongs to the types *verb*, *transitive verb*, *present-tense verb* and *third-person verb*. In an inheritance hierarchy, a type B inherits from (is a subtype of) another type A, if and only if the set of feature structures described by B is a subset of the set of feature structures described by A. The inheritance hierarchies of SBCG are referred to as *multiple inheritance hierarchies* because a given type can inherit properties from multiple dominating types (e.g., *present tense* and *transitive*).

Crucially for our purposes, SBCG generalizes the lexical-inheritance model as the appropriate model for the relations among constructions. The rationale is that, as observed by Jackendoff (1997, Chapter 7) and Croft and Cruse (2002, Chapter 9), constructions mean what they mean in the same way that words do. Like words, constructions can invoke semantic, pragmatic, and phonological conditions simultaneously. As an example of an idiomatic pattern with highly particular intonational phonology, consider the exclamatory construction that Michaelis and Lambrecht (1996) refer to as the Antitopic Exclamative. In this construction, a preclausal interjection receives prosodic prominence and the following clause receives the intonational contour of a right-dislocated phrase. Examples of the Antitopic Exclamative are given in (12–14):

(12) GOD it's hot.

(13) MAN that's loud.

(14) DAMN I'm good.

The point here is that, as Croft and Cruse (2002: 247) put it, "[c]onstructions, like the lexical items in the lexicon, are 'vertical' structures that combine syntactic, semantic and even phonological information (for the specific words in a construction, as well as any unique prosodic features that may be associated with a construction". The more general point, as expressed by Culicover and Jackendoff (2005: 15) is that there is "a continuum of grammatical phenomena from idiosyncratic (including words) to general rules of grammar".

As an example of an inheritance hierarchy for constructions, consider the following functions of the pattern commonly referred to as subject–auxiliary inversion (SAI), taken from Fillmore (1999):

(15) Yes-no question: Has he left?

(16) Inverted exclamation: Am I tired!

(17) Negative adverb preposing: Never will I harm you.

(18) Information question: What would you do?

(19) Optative: May it be so!

(20) Conditional: Had we been there, we could have stopped him.

In SBCG, as described by Sag (2007), the auxiliary-initial clausal pattern is a type (of construct), and various constructions, like those exemplified above, mention this type in their consequent clauses. For example, the exclamative SAI construction illustrated in (16), has *inverted–exclamative–construct* (*inv–excl–cxt*) as its antecedent, while its consequent invokes the more general construction *auxiliary–initial–construct* (*ai–cxt*), as illustrated in (21):

(21) $inv-excl-cxt \Rightarrow \begin{bmatrix} ai-cxt \\ \ldots \end{bmatrix}$

In (21), the type to which the inverted exclamative belongs is represented by the label *ai–cxt* at the top of the feature matrix; this label represents the sign type. The additional features required to capture the properties unique to the inverted exclamative are not mentioned here, and are represented by ellipses [...]. The property common to all of the constructions in (21) is the use of an auxiliary-initial clause, but each construction also has idiosyncratic properties; for example, (17) requires a negative adverb in clause-initial position. In addition, each construction has an idiosyncratic communicative function (e.g., requesting information, exclaiming). These are functions that one would not know simply by knowing that a given construction is an instance of the SAI pattern.

7.3 Constructions license arguments and complements

7.3.1 Constructions as a source of valence variability

Where does a verb's frame come from? The obvious answer is the verb itself, and this is the answer that syntacticians have traditionally provided, whether they view predicate–argument relations as syntactic sisterhood (as per constituent-structure-based models) or as a lexical property (the verb's combinatoric potential). Thus, Haegeman, in her introduction to Government and Binding theory, states, "the thematic structure of a predicate, encoded in the theta grid, will determine the minimal components of the sentence" (Haegeman 1994: 55). Similarly, Bresnan, in her introduction to Lexical-Functional Grammar, states, "On the semantic side, argument structure represents the core participants events (states, processes) designated by a single predicator. [...] On the syntactic side, argument structure represents the minimal information needed to characterize the syntactic dependents of an argument-taking head" (Bresnan 2001: 304). It is difficult, however, to square this view with the observation, made by Goldberg (1995; 2006), Partee and Borschev (2007), and Michaelis and Ruppenhofer (2001), among others, that a verb can often be found in unexpected frames, which nonetheless make sense in context. For example, as shown in (22–24), single-argument activity verbs like *melt* and *sparkle*, which have nothing intrinsically to do with location, can appear in the "locative-inversion" pattern:

(22) In Maria's sticky hand *melted* a chocolate-chip ice-cream cone. (Birner and Ward 1998: 193)

(23) And in this lacy leafage *fluttered* a number of grey birds with black and white stripes and long tails. (Levin and Rappaport Hovav 1995: 226)

(24) Down at the harbor there is teal-green clubhouse for socializing and parties. Beside it *sparkles* the community pool. (*Vanity Fair*, 8/01)

In (22–24), the verb appears to describe what an entity is doing while in its location (melting, fluttering, sparkling) rather than a location state *per se*. Looking at a similar class of examples in Russian, Partee and Borschev (2007: 158) observe, "[o]ne could say that THING and LOC are roles of the verb [*be*], but it is undoubtedly better to consider them roles of the participants of the situation (or state) of existing or of being located". If one were to alter the preceding quote by replacing the words *situation (or state) of existing or of being located* with the words *locative-inversion construction*, it would express the constructional view of verbal argument-structure, first articulated by Goldberg (1995; 2002; 2006). Goldberg argues that argument-structure patterns are constructions that denote situation types and that a verb's meaning and combinatory potential can change to fit the meaning of a given construction (see also Michaelis and Ruppenhofer 2001 and Michaelis 2004). The construction-based model of argument structure proposed by Goldberg is based on the idea that verb meanings are combined with construction meanings via a fixed number of semantic relations (including *instance*, *means*, and *manner*) and the semantic-role list licensed by the verb may accordingly be augmented up to that licensed by the construction. Examples are given in (25–26):

(25) Most likely they were fellow visitors, just *panting* up to the sky-high altar out of curiosity. (L. Davis, *Last Act in Palmyra*, p. 28)

(26) When a visitor passes through the village, young lamas stop picking up trash to mug for the camera. A gruff "police monk" *barks* them back to work. (*Newsweek* 10/13/97)

In (25), *pant*, a verb that otherwise licenses only a single argument, appears with two: it denotes the *manner* of the directed-motion event denoted by the construction. In (26), *bark*, another otherwise monovalent activity verb, has two additional arguments, a direct object and an oblique expression that indicates direction; in this context, the verb denotes the *means* by which a metaphorical caused-motion event, denoted by the construction, occurs. Rather than presuming a nonce lexical entry for *pant* in which it means "move toward a goal while panting" and for *bark* in which it means "move something from one place to another by barking", a constructionist presumes that the verbs in (25–26) mean what they always mean; arguments not licensed by the verb are licensed by the construction with which the verb combines. The constructional model of verbal syntactic variability is therefore more parsimonious than a lexicalist one: by using a small number of argument-structure constructions, it limits the number of lexical entries needed for each verb.

7.3.2 Weird sisterhood

A number of argument-structure patterns involve verbal complementation patterns that are not licensed by the general-purpose head-complement or specifier-head phrase-building rule schemas. Many of these patterns have specialized communicative functions. A look at these phenomena suggests that fine-grained construction, rather than non-category-specific phrase-structure rules, pair predicates and their complements. In this subsection, we will look at three cases of weird sisterhood found in English: Nominal Extraposition, *Just because*, and Hypotactic Apposition. The data are taken from one of two corpora of English telephone conversations that are available through the Linguistic Data Consortium (www.ldc.upenn.edu): the Switchboard corpus and the Fisher corpus.

7.3.2.1 *Nominal extraposition*

In nominal extraposition, an exclamatory adjective, e.g., *amazing*, licenses an NP complement:

(27) I know it's just it's unbelievable the different things that are happening in America today.

(28) I know. I love that game. It's amazing the words they come up with.

The pattern exemplified in (27–28) is idiosyncratic in two respects. First, adjectives are not case-assigning predicators and should not therefore license direct objects. Second, this NP complement is interpreted as denoting a scalar degree (Michaelis and Lambrecht 1996). In (28), for example, the NP *the words they come up* stands in for a scalar expression like "the number of words they come up with". The fact that the complement of *amazing* in (28) has a scalar interpretation follows from the fact that (28) is an exclamation, but the pairing of an exclamatory adjective with an NP sister that denotes a degree, metonymically or otherwise, requires a construction that provides for this syntax and this meaning.

7.3.2.2 *Just because*

In the *Just Because* construction, a negated epistemic verb, typically *mean*, licenses a finite clause subject introduced by *just because* (Bender and Kathol to appear):

(29) Just because they use primitive means of doing things does not mean that they can't expand.

(30) Just because they say it doesn't mean that's the only way to look at it.

Clausal subjects are ordinarily introduced by the complementizer *that*, not by a subordinating conjunction. For this reason, one cannot use the phrase-structure rule that pairs a specifier with a head to account for the pattern illustrated in (29–30). Instead, as Bender and Kathol argue, the grammar of English must contain an argument-structure construction that allows the verb *mean*, when negated, to license a clausal subject introduced by *just because*.

7.3.2.3 *Hypotactic Apposition*

When English speakers announce forthcoming propositional content using a cataphoric demonstrative pronoun, they may do so by means of either the paratactic construction exemplified in (31) or the subordinating construction illustrated in (32–33), in which the asserted proposition is expressed by a clausal complement of the copula:

(31) Yeah, well, that's another problem: I think to really correct the judicial system you have to get the lawyers out of it.

(32) That's the problem is that they just hate us so much and I never re- I never really realized.

(33) That's the main thing is that I can't tell whether the thing is going to fit.

Sentence (33) is an example of the construction that Brenier and Michaelis (2005) refer to as Hypotactic Apposition. In Hypotactic Apposition, the verb *be* combines with two arguments that it would not ordinarily: a clause containing the pronoun *that* (in (32), e.g., *that's the problem*) and a clausal complement to which this *that* refers (in (32), *they just hate us so much*). This is not the ordinarily combinatoric behavior of equational *be*, as illustrated in (34):

(34) The problem is that they just hate us so much.

In (34), *be*, as expected, combines with a subject NP and a clause. Thus, the combinatoric behavior of *be* in (32–33) cannot be attributed to the lexeme *be* but must instead be attributed to the Hypotactic Apposition construction.

7.3.3 Lexical-class constructions

The constructions needed to account for valence augmentation and weird sisterhood have not yet been described. The constructions in question are referred to in SBCG as *lexical-class constructions*. Lexical-class constructions describe the properties of a class of lexical entries. These properties include but are not limited to: the semantic frame of the lexeme, the syntactic category of the lexeme's semantic roles and contextual attributes like illocutionary force. Lexical-class constructions have the general form shown in (35), where *lex* stands for any subtype of the type *lexeme*:

(35) lex ⇒ [. . .]

As Sag (2007) argues, there is no formal difference between lexical-class constructions and those that describe constructs (the latter of which Sag 2007 refers to as *combinatoric constructions*). The only difference is the nature of the type name that serves as the antecedent of the constraint. How could it be that a construction that describes a mother–daughter combination could be the same as one that describes a word or lexeme? The answer is that both lexical-class and combinatoric constructions describe signs. In the case of a combinatoric construction this sign happens to have a DTRS feature, ensuring that it can license a mother node in a local tree in a construct, but this mother node is a sign like any other. Crucially, lexical-class constructions can combine with one another, creating a highly specific lexeme entry. Among lexical-class constructions are those that allow a required semantic role of the verb to be missing. These constructions are referred to as null-instantiation constructions (Fillmore et al. in prep.). Null-instantiation constructions eliminate a semantic role from the verb's valence list while ensuring (through the addition of a quantifier frame to the verb's FRAMES list) that the missing valence member is interpreted as an existentially or anaphorically bound variable (Fillmore et al. in prep.). An example of null instantiation is given in (36):

(36) I cried into my beer [when I saw this story about walruses appearing on the Alaskan coast]. (Google)

When we interpret (36), we understand that there is some entity (lachrymal fluid) that the speaker caused to move into the beer, but no such entity is expressed in the sentence. While (36) expresses a caused-motion event akin to the *bark* sentence (26), the theme argument is missing. This is not of course a special fact about (36), since the theme argument of *cry* is not generally expressed: *I cried (many tears) during that movie*. Goldberg (2005) proposes a null-instantiation construction for verbs of emission like *cry*, *spit*, and *bleed*. This construction allows such verbs to appear without their theme arguments. Examples like (36) are produced by the interaction of the caused-motion and the emission-verb lexical-class constructions.

7.4 There is a continuum of idiomaticity

As foundation of construction-based syntax is the idea that rules of syntactic combination are directly associated with interpretive and use conditions, in the form of semantic and pragmatic features that attach to the mother or daughter nodes in these descriptions (Sag 2007; 2008). This amounts to the claim that syntactic rules mean things. Meaning, of course, is generally viewed as the exclusive purview of words, and in the prevailing view of meaning composition, syntactic rules do no more than determine what symbol sequences function as units for syntactic purposes. So, while syntactic rules assemble words and their dependent elements into phrases, and the phrases denote complex concepts like predicates and propositions, the rules cannot add conceptual content to that contributed by the words; nor can they alter the combinatoric properties of the words. On this view, which Jackendoff (1997: 48) describes as the *doctrine of syntactically transparent composition*, "[a]ll elements of content in the meaning of a sentence are found in the lexical conceptual structures [...] of the lexical items composing the sentence" and "pragmatics plays no role in determining how [lexical conceptual structures] are combined".

To embrace a construction-based model of semantic composition is not to reject the existence of syntactically transparent composition but instead to treat it, as Jackendoff recommends (1997: 49), as a "default in a wider array of options". That is, whenever a class of expressions can be viewed as licensed by a context-free phrase structure rule accompanied by a rule composing the semantics of the mother from the semantics of the daughters, a construction-based approach would propose a construction that is functionally identical to such a form–meaning pairing (Kay and Michaelis in press). But constructional approaches also provide a way to represent linguistic structures whose meanings are more than the sum of their parts. A case in point is the negative polar question. An affirmative question like (37) requests an evenhanded evaluation of its propositional content, expressed in (38):

(37) Did the Magna Carta change the way the king behaved?

(38) The Magna Carta changed the way the king behaved.

However, a negative question like (39) is not understood as posing a negative proposition and requesting an evenhanded evaluation of its truth or falsity, as in (40). Instead, the negative question, like the tag question in (41), suggests that the *affirmative* proposition is true:

(39) Didn't the Magna Carta change the way the king behaved?

(40) True or false: The Magna Carta didn't change the way the king behaved.

(41) The Magna Carta changed the way the king behaved, didn't it?

A constructional approach allows the grammar to capture the straightforward cases of compositionality and also the cases, like negative questions, in which the construction adds to or otherwise changes what simple compositionality might predict. A further illustration of a construction that is syntactically regular and yet semantically opaque is provided by the WXDY construction,[1] exemplified in (42):

(42) What's this fly doing in my soup?

The ambiguity of (42) is known to anyone familiar with the old joke in which it serves as the set-up (eliciting the response *the backstroke* from an obtuse waiter). Kay and Fillmore (1999) argue that the sentence pattern in (42) has both a transparent interpretation (in which it inquires about someone's actions) and an idiomatic interpretation, in which it is a *why* questions used to inquire about a situation that strikes the speaker as anomalous. Kay and Fillmore posit a WH-question construction, WXDY, to which the latter interpretation attaches. Among other formal conditions, WXDY fixes the interrogative word as *what* and requires the form of the main verb to be progressive. In WXDY, as in the case of the negative question in (39) and the exclamatory pattern described in 7.3.2.1, an illocutionary force attaches to a clause pattern rather than to any particular word in that pattern. What this means in SBCG terms is that illocutionary force belongs to the contextual features in such constructions' mother signs (Sag 2008; Ginzburg and Sag 2000).

At the subclausal level, there are many idiomatic constructions that create similarly ambiguous word strings. For example, (43) may mean what it means either because it instantiates an idiomatic VP construction (whose meaning is "jokingly mislead") or because it instantiates the more general constructions that combine nominals with possessive determiners, auxiliary verbs with their complements, and NPs with VPs:

(43) She's pulling my leg.

Under strictly syntactic composition, the ambiguous (43) would require two different syntactic representations. This is an undesirable result, because the two meanings of (43) cannot be attributed to a bracketing ambiguity like that in (44):

(44) She saw her neighbor with a telescope.

Under a constructional approach, the two meanings of (44) are described by two different collections of constructions. But construction-based composition is still rule-based: an interpreter who knows all of the words, and *all* of the rules that combine words and phrases into larger units, also knows the forms and meanings of all the larger units, including all the sentences. Constructional approaches focus on the fact that there are a great many rules, and that many of these rules attach semantic interpretations directly to complex syntactic objects.

[1] The WXDY construction is so called because it consists of the question word *why* followed by (in order) an NP of X type (the subject of the inverted progressive auxiliary *be*), the gerundial verb *do*, and a PP or other secondary predicate (represented by the variable Y).

7.5 Core and periphery are interleaved during production

As described in section 7.4, the construction grammarian conceives of a language as presenting a continuum of idiomaticity, or generality, of expressions; a construction grammar models this continuum with an array of constructions of correspondingly graded generality (Fillmore et al. 1988). Inheritance networks capture the relationships that exist between general constructions of potentially universal significance, like coordination, and inarguably language-particular patterns like the adverbial expression *by and large*—perhaps the only coordinate structure in English that features a conjoined preposition and adjective. But construction grammarians see no obvious point along the continuum from schema to formula where one can draw the line between "core" and "periphery". It seems common practice to include in the core both the obvious cases and as much of the rest of the language as fits the theoretical apparatus at hand (Culicover and Jackendoff 1999). But the resulting models cannot then be portrayed as theories "of the core" without circularity. Evidence for the inseparability of core and periphery comes from the interleaving of the two during production: stretches of speech licensed by idiomatic constructions can contain within them stretches licensed by "regular rules" and *vice versa*. This is illustrated in (45):

(45) Me try to pull the leg of a philosopher? No way.

Sentence (45) illustrates the Incredulity Response construction, which, according to Lambrecht (1990), consists of (a) a property predicate (e.g., *pull the leg of a philosopher*), (b) an entity (e.g., *me*), and (c) an expression of incredulity concerning the entity's membership in the class of individuals named by the property predicate. Formally, the entity is expressed by an accusative-case nominal and the predicate by a nonfinite VP or other phrase. Lambrecht argues that the Incredulity Response is a topic-comment construction, and that the entity and predicate are detached topics. Evidence for the latter claim comes from the fact that the two constitute distinct intonation units and can be reordered with respect to one another (as in *Pull the leg of a philosopher? Me?*). While this construction performs a basic communicative function—commenting on the validity of someone's prior assertion—it does so in a way that owes little or nothing to the ordinary English syntax of predication and subordination. It is equally obvious, however, that both the entity constituent and the predicate constituent are licensed by regular or "core" constructions of English—only their combination in the Incredulity Response construction is idiomatic. Moreover, coterminous with the syntactically transparent VP *try to pull the leg of a philosopher*, we find the VP idiom *pull the leg of a philosopher*, licensed by the idiomatic *pull-someone's-leg* construction, and going further inside the NP *the-leg-of-a-philosopher*, which is licensed by the idiomatic *pull-someone's-leg* construction, we find the transparent genitive PP *of a philosopher*.

Thus, it is unlikely that grammar consists of a set of productive rules, a lexicon, and a collection of frozen phrasal idioms. Instead, these "modules" appear to be permeable.

7.6 Constructions have properties that do not map

An advantage of modeling constructions in a multiple-inheritance hierarchy is that it provides a succinct way of describing the relations among families of similar constructions, indicating which properties they share and which are peculiar to each maximal (or leaf) construction (that is, each construction that has no sub-constructions). The family of SAI constructions discussed in section 7.2 above provides an illustration. While the family of SAI constructions represents a one-to-many form-function mapping, inheritance hierarchies are also used to describe many-to-one form-function mappings, as in Michaelis and Lambrecht's (1996) study of English exclamatory constructions. They analyze a range of English exclamations—including the bare NP exclamative illustrated in (46), nominal extraposition, as described in section 7.3.2.1 above, and subordinate-clause exclamations, as in (47):

(46) The nerve!

(47) I can't believe the nerve of some people.

They capture the shared interpretive and use constraints on these patterns by treating each exclamative sentence type as an instance of an abstract exclamatory construction, whose semantico-pragmatic features include scalar meaning, a specific epistemic stance of the speaker, and property attribution. Thus, relations of family resemblance are posited both on formal and semantic grounds.

Of course, one might observe that what inheritance networks do is something that procedural approaches have long done: represent those situations in which two different verb frames or syntactic tree structures share a single event-structure representation, as in the transformational accounts of passive, topicalization, and raising. Certainly, inheritance networks provide a declarative rather than procedural mechanism for describing this shared structure, but one could legitimately ask whether the type hierarchy of SBCG is a mere notational variant of the familiar lexical and syntactic mapping procedures. The answer is no, for two reasons.

The first reason is that procedural approaches to argument-structure variability, unlike declarative approaches, presuppose bilateral entailment relationships between argument-structure affordances, as a conceptual necessity: if two verbal argument-structures are to be mediated by a rule, the existence of frame A for a given verb entails the existence of frame B for that verb, and vice versa. For example, if a verb takes a direct object, it should also be found in the passive voice. But as scholars ranging from Lakoff (1970) to Pinker (1989) have observed, rules have abundant lexical exceptions. One could argue that this fact lowers the level of generality that a procedural approach is supposed to achieve. And while this objection would be fair, the objection made by constructionists (e.g., Goldberg 1995; 2002; Michaelis and Ruppenhofer 2001) is actually stronger: there are in fact two classes of lexical exceptions, and only one of these is countenanced by the procedural approach. In the first class are those verbs that fail to undergo a given rule. For example, Latinate verbs like *contribute* do not allow "dative movement": **She contributed the campaign a donation*. Pinker (1989) suggests that such exceptions are principled and proposes that certain semantically defined lexical classes block the application of lexical rules. This is certainly a more stipulative approach than one might seek in a grammar based on abstract constraints, but it does increase descriptive adequacy. The second class is more troublesome. It includes "output" patterns that lack the requisite input structure. The existence of such examples suggests that the derivational approach to verb-valence variability is not the right model of this domain. Examples from this second class are given in (48–50):

(48) Ditransitive (double-object): She begrudges/envies me my success. (cf. *She begrudges/envies my success to me.)

(49) Raising: She seems/appears to have left. (cf. *That she has left appears/seems.)

(50) Passive: She is said/rumored to have won. (cf. *They said/rumored her to have won.)

In each of these examples, we see that the putative input structure is ungrammatical, whether it is the oblique-goal frame in (48), the clausal-subject frame in (49), or the active-voice frame in (50). The essential observation is that the lexemes in question (e.g., *begrudge*, *seem*, *say*) lack one of the two argument-structure frames that procedural approaches place into correspondence (lexical or transformational). As we have seen, it is possible to block a verb from undergoing a rule, but if ditransitive, raising, and passive lexemes (or trees) are the *products* of rules, the procedural approaches incorrectly predict the ungrammaticality of (48–50). In SCBG, by contrast, invariant lexeme entries, like that of *begrudge*, are represented as more fully specified than those of variable lexeme entries, like *give*. Most typically, the additional specification takes the form of a CASE attribute attached to one of the verb's valence members. Because SBCG is unification-based, the additional feature prevents the entry from combining with combinatoric and lexical-class constructions that contain conflicting specifications (Sag 2007).

Let us now turn to the second reason that the construction-based approach to argument structure is distinct from one that uses (syntactic or semantic) procedures to alter verb entries: the constructional approach captures semantic and pragmatic conditions that may be unique to each of the two putative structural alternates (Goldberg 1995). Mapping procedures, if they are to operate compositionally, cannot introduce meanings into the output structure. However, as observed by Goldberg (1995, Chapter 5) and Michaelis and Ruppenhofer (2001), a wide array of verb frames held to be the outputs of lexical rules have entailments that they do not share with their input frames. These entailments include the requirement that the goal argument of a ditransitive verb be interpreted as a recipient and that the location argument of an applicative (*be*-prefixed) verb in German be construed as a surface. Because constructions can have as many specialized communicative and interpretive conditions as words do, such idiosyncrasies are easy to describe if the two verb frames (e.g., ditransitive and oblique goal) are taxonomic sisters in an inheritance hierarchy (Sag 2007).

7.7 Conclusion

The focus of syntactic theory has long been on determining the range of possible human languages—a goal that for Chomsky (1995: 435) and adherents justifies a reduction in the range of linguistic facts that the theory should seek to cover. Construction grammarians retain a commitment to thorough description of individual language grammars. It might therefore appear that they lack interest in explanatory theories of grammar and seek only to demonstrate the infinite diversity of language. In fact, SBCG makes strong universal claims, including the Sign Principle and the Head-Feature Principle (Sag 2007; 2008). But theory comparison in this arena is hindered by the fact that many potential universals cannot be disentangled from the formal conventions of particular theories. This seems particularly true of universals assumed by proponents of the so-called Principles and Parameter model, as in the following quote:

> The history of syntactic investigation is marked by a small number of central discoveries which created the syntactician's research agenda. One can divide these discoveries into two groups: the discovery of hierarchical constituent structure, and the discovery that elements may occupy more than one position within this hierarchy, which the literature calls movement. (Pesetsky 1997: 134)

To view "movement" as a "discovery" is to confuse representational conventions with linguistic facts. It is illogical to view construction-based syntax as anti-universalist because it does not assume a universal grammar based on such conventions. The two putative discoveries referred to above are in fact simply mutually reinforcing assumptions. The need to capture relationships between constructions by relating them transformationally comes from the assumption that syntax is autonomous, which in turn requires that semantic features play no role in syntactic generalizations. The result is that the syntactician cannot relate two constructions by describing them as alternate syntactic realizations of a given semantic role; she or he must instead speak of procedures that change the position of a given syntactic constituent in hierarchical syntactic structure. And of course transformations are what make it possible to maintain that all languages have hierarchical constituent structure (and that this structure underlies the assignment of morphological case, among other things): in free-word order languages, the lack of observable constituent structure is attributed to permutations called "scrambling".

Because the circularity of the Chomskyan principles makes them virtually immune to falsification, constructionists have aimed instead at the other *major* foundation of Chomskyan universal grammar: language-particular parameter settings. Pullum and Zwicky (1991) argue, for example, that the prohibition against double-*ing* sequences in English (e.g., **stopping walking*) is not a "transconstructional filter" but a constraint on a single constituency-defining rule. And Van Valin and LaPolla (1997, Chapter 6) have shown that the patterns of semantic neutralization and restriction that define syntactically privileged arguments (e.g., subjects) vary not merely from language to language but from construction to construction *within* a given language. An illustration is found in English adjectival secondary predicates that denote a resultant state:

(51) She hammered the metal flat.

While one might assume that the entity undergoing the change of state in such sentences is appropriately described as the direct object, this would be an incorrect assessment because that entity can also be expressed by a subject NP:

(52) The cake fell flat.

What unites the changed entities in (51–52) is that both are patient arguments. This suggests that the English construction that licenses secondary predicates of result semantically features the pattern of semantic-role restriction characteristic of ergative-absolutive case systems. What might otherwise be said to characterize a language (e.g., the nominative-accusative or ergative-absolutive pattern of neutralization) in fact characterizes a construction. Phenomena that have been taken as evidence of nominative-accusative or ergative-absolutive "undercurrents" in a given language are more accurately viewed as effects of construction-particular argument-selection patterns. Such phenomena therefore need not be taken as evidence of instability in a grammatical system, since they are natural consequences of construction-particular constraints. Syntactic generalizations may not be nearly as general as we have come to believe.

CHAPTER 8

CORPUS-BASED AND CORPUS-DRIVEN ANALYSES OF LANGUAGE VARIATION AND USE

DOUGLAS BIBER

8.1 Introduction

Corpus linguistics is a research approach that has developed over the past several decades to support empirical investigations of language variation and use, resulting in research findings that have much greater generalizability and validity than would otherwise be feasible. Corpus linguistics is not in itself a model of language. In fact, at one level it can be regarded as primarily a methodological approach:

- it is empirical, analyzing the actual patterns of use in natural texts;
- it utilizes a large and principled collection of natural texts, known as a "corpus", as the basis for analysis;

- it makes extensive use of computers for analysis, using both automatic and interactive techniques;
- it depends on both quantitative and qualitative analytical techniques (Biber et al. 1998: 4).

At the same time, corpus linguistics is much more than a methodological approach: these methodological innovations have enabled researchers to ask fundamentally different kinds of research questions, sometimes resulting in radically different perspectives on language variation and use from those taken in previous research. Corpus linguistic research offers strong support for the view that language variation is systematic and can be described using empirical, quantitative methods. Variation often involves complex patterns consisting of the interaction among several different linguistic parameters, but, in the end, it is systematic. Beyond this, the major contribution of corpus linguistics is to document the existence of linguistic constructs that are not recognized by current linguistic theories. Research of this type—referred to as a "corpus-driven" approach—identifies strong tendencies for words and grammatical constructions to pattern together in particular ways, while other theoretically possible combinations rarely occur. Corpus-driven research has shown that these tendencies are much stronger and more pervasive than previously suspected and that they usually have semantic or functional associations (see section 8.3 below).

In some ways, corpus research can be seen as a logical extension of quantitative research in sociolinguistics begun in the 1960s (e.g., Labov 1966), which rejected "free variation" as an adequate account of linguistic choice and argued instead for the existence of linguistic variable rules (see Chambers and Trudgill 1980: 59–61; 146–9). However, research in corpus linguistics differs from quantitative sociolinguistic research in at least two major ways:

(1) Quantitative sociolinguistics has focused on a relatively small range of varieties: usually the social dialects that exist within a single city, with secondary attention given to the set of "styles" that occur during a sociolinguistic interview. In contrast, corpus research has investigated the patterns of variation among a much wider range of varieties, including spoken and written registers as well as dialects.

Corpus-based dialect studies have investigated national varieties, regional dialects within a country, and social dialects. However, the biggest difference from quantitative sociolinguistics here has to do with the investigation of situationally-defined varieties: "registers". Quantitative sociolinguistics has restricted itself to the investigation of only spoken varieties, and considered only a few "styles", which speakers produce during the course of a sociolinguistic interview (e.g., telling a story vs. reading a word list). In contrast, corpus-based research investigates the patterns of variation among the full set of spoken and written registers in a language. In speech, these include casual face-to-face conversation, service encounters, lectures, sermons, political debates, etc.; and, in writing, these include e-mail messages, text-messaging, newspaper editorials, academic research articles, etc.

(2) Quantitative sociolinguistics has focused on analysis of "linguistic variables", defined such that the variants must have identical referential meaning. Related to this restriction, quantitative sociolinguistic research has focused exclusively on non-functional variation. For these reasons, most quantitative sociolinguistic research has focused on phonological variables, such as [t] vs. [θ]. Sociolinguistic variation is described as indexing different social varieties, but there is no possibility of functional explanations for why a particular linguistic variant would be preferred in one variety over another.

In contrast, corpus research considers all aspects of language variation and choice, including the choice among roughly synonymous words (e.g., *big, large, great*), and the choice among related grammatical constructions (e.g., active vs. passive voice, dative movement, particle movement with phrasal verbs, extraposed vs. subject complement clauses). Corpus-based research goes even further, investigating distributional differences in the extent to which varieties rely on core grammatical features (e.g., the relative frequency of nouns, verbs, prepositional phrases, etc.). All of these aspects of linguistic variation are interpreted in functional terms, attempting to explain the linguistic patterns by reference to communicative and situational differences among the varieties. In fact, much corpus-based research is based on the premise that language variation is functional: that we choose to use particular linguistic features because those forms fit the communicative context of the text, whether in conversation, a political speech, a newspaper editorial, or an academic research article.

In both of these regards, corpus-based research is actually more similar to research in functional linguistics than research in quantitative sociolinguistics. By studying linguistic variation in naturally occurring discourse, functional linguists have been able to identify systematic differences in the use of linguistic variants. An early study of this type is Prince (1978), who compares the distribution and discourse functions of WH-clefts and *it*-clefts in spoken and written texts. Thompson and Schiffrin have carried out numerous studies in this research tradition: Thompson on detached participial clauses (1983), adverbial purpose clauses (1985), omission of the complementizer *that* (Thompson and Mulac 1991*a*; 1991*b*), relative clauses (Fox and Thompson 1990); and Schiffrin on verb tense (1981), causal sequences (1985*a*), and discourse markers (1985*b*). Other early studies of this type include Ward (1990) on VP preposing, Collins (1995) on dative alternation, and Myhill (1995; 1997) on modal verbs.

More recently, researchers on discourse and grammar have begun to use the tools and techniques available from corpus linguistics, with its greater emphasis on the representativeness of the language sample, and its computational tools for investigating distributional patterns across registers and across discourse contexts in large text collections (see Biber et al. 1998; Kennedy 1998; Meyer 2002; and McEnery et al. 2006). There are a number of book-length treatments reporting corpus-based investigations of grammar and discourse: for example, Tottie (1991*a*) on negation, Collins (1991) on clefts, Mair (1990) on infinitival complement clauses, Meyer (1992) on apposition, Mindt 1995 on modal verbs, Hunston and Francis (2000) on pattern grammar, Aijmer (2002) on discourse particles, Rohdenburg and Mondorf (2003) on grammatical variation; Lindquist and Mair (2004) on grammaticalization, Mahlberg (2005) on general nouns, Römer (2005) on progressives.

A central concern for corpus-based studies is the representativeness of the corpus (see Biber 1993; Biber et al. 1998: 246–50; McEnery et al. 2006: 13–21, 125–30). Two considerations are crucial for corpus design: size and composition. First, corpora need to be large enough to accurately represent the distribution of linguistic features. Second, the texts in a corpus must be deliberately sampled to represent the registers in the target domain of use.

Corpus studies have used two major research approaches: "corpus-based" and "corpus-driven". Corpus-based research assumes the validity of linguistic forms and structures derived from linguistic theory; the primary goal of research is to analyze the systematic patterns of variation and use for those pre-defined linguistic features. One of the major general findings from corpus-based research is that descriptions of grammatical variation and use are usually not valid for the language as a whole. Rather, characteristics of the textual environment interact with register differences, so that strong patterns in one register often represent weak patterns in other registers. As a result, most corpus-based studies of grammatical variation include consideration of register differences. The recent *Longman Grammar of Spoken and Written English* (Biber et al. 1999) is the most comprehensive reference work of this kind, applying corpus-based analyses to show how any grammatical feature can be described for its patterns of use across discourse contexts and across spoken and written registers.

In contrast, "corpus-driven" research is more inductive, so that the linguistic constructs themselves emerge from analysis of a corpus. The availability of very large, representative corpora, combined with computational tools for analysis, make it possible to approach linguistic variation from this radically different perspective. The corpus-driven approach differs from the standard practice of linguistics in that it makes minimal a priori assumptions regarding the linguistic features that should be employed for the corpus analysis. In its most basic form, corpus-driven analysis assumes only the existence of words, while concepts like "phrase" and "clause" have no a priori status. Rather, co-occurrence patterns among words, discovered from the corpus analysis, are the basis for subsequent linguistic descriptions.

The following sections illustrate the kinds of analyses and perspectives on language use possible from both corpus-based and corpus-driven approaches.

Section 8.2 illustrates the corpus-based approach, which documents the systematic patterns of language use, often showing that intuitions about use are wrong. Section 8.3 then illustrates the corpus-driven approach, showing how corpus research can uncover linguistic units that are not detectable using the standard methods of linguistic analysis.

8.2 Corpus-based research studies

As noted above, the corpus-based approach has some of the same basic goals as research in functional linguistics generally, to describe and explain linguistic patterns of variation and use. The goal is not to discover new linguistic features but rather to discover the systematic patterns of use that govern the linguistic features recognized by standard linguistic theory.

One major contribution of the corpus-based approach is that it establishes the centrality of register for descriptions of language use. That is, corpus-based research has shown that almost any linguistic feature or variant is distributed and used in dramatically different ways across different registers. Taken together, corpus-based studies challenge the utility of general linguistic descriptions of a language; rather, these studies have shown that any linguistic description that disregards register is incomplete or sometimes even misleading.

Considered within the larger context of quantitative social science research, the major strengths of the corpus-based approach are its high reliability and external validity. The use of computational tools ensures high reliability, since a computer program should make the same analytical decision every time it encounters the same linguistic phenomenon. More importantly, the corpus itself is deliberately constructed and evaluated for the extent to which it represents the target domain (e.g., a register or dialect). Thus, the linguistic patterns of use described in corpus-based analysis are generalizable, explicitly addressing issues of external validity.

However, judged by the normal interests of linguists, the greater contribution of the corpus-based approach is that it often produces surprising findings that run directly counter to our prior intuitions. That is, as linguists we often have strong intuitions about language use (in addition to intuitions about grammaticality), believing that we have a good sense of what is normal in discourse. While it is difficult to evaluate intuitions about grammaticality, intuitions about use are open to empirical investigation. Corpus-based research is ideally suited for this task, since one of the main research goals of this approach is to empirically identify the linguistic patterns that are extremely frequent or rare in discourse from a particular variety. And when such empirical investigations are conducted, they often reveal patterns that are directly counter to our prior expectations.

A simple case study of this type, taken from the *Longman Grammar of Spoken and Written English* (Biber et al. 1999: 460–3), concerns the distribution of verb aspect in English conversation. There are three aspects distinguished in English verb phrases:

Simple aspect:	*Do you like it?*
Progressive aspect:	*I was running around the house like a maniac.*
Perfect aspect:	*You haven't even gone yet.*

The question to consider is which grammatical aspect is most common in face-to-face conversation?

It is much easier to illustrate the unreliability of intuitions in a spoken lecture because audience members can be forced to commit to an answer before seeing the corpus findings. For full effect, the reader here should concretely decide on an answer before reading further.

Hundreds of linguists have been polled on this question, and the overwhelming majority have selected progressive aspect as the most common verb aspect in English conversation. In fact, as Figure 8.1 shows, progressive aspect is more common in conversation than in other registers. The contrast with academic prose is especially noteworthy: progressive aspect is rare in academic prose but common in conversation.

However, as Figure 8.2 shows, it is not at all correct to conclude that progressive aspect is the most common choice in conversation. Rather, simple aspect is clearly the unmarked choice. In fact, simple aspect verb phrases are more than 20 times as common as progressives in conversation.

The following conversation illustrates this extreme reliance on simple aspect (underlined) in contrast to the much more specialized use of progressive aspect (in ***bold italics***):

Jan Well girls we better open the presents, I'm going to fall asleep.
Kris I know.
Amanda Okay, right after he rolls out this last batch.
Rita Your face is really hot. Why ***are*** you ***leaving*** it, we'***re*** not ***leaving*** till Sunday are we?
Jan Which ever day you prefer, Saturday or Sunday.
Rita When ***are*** you ***leaving***?
Amanda Sunday morning.
Rita Oh, well we don't have to do it right away.
Kris Oh well let's just do it.
Rita I'd rather wait till I feel like it.
Jan But we'***re doing*** it.

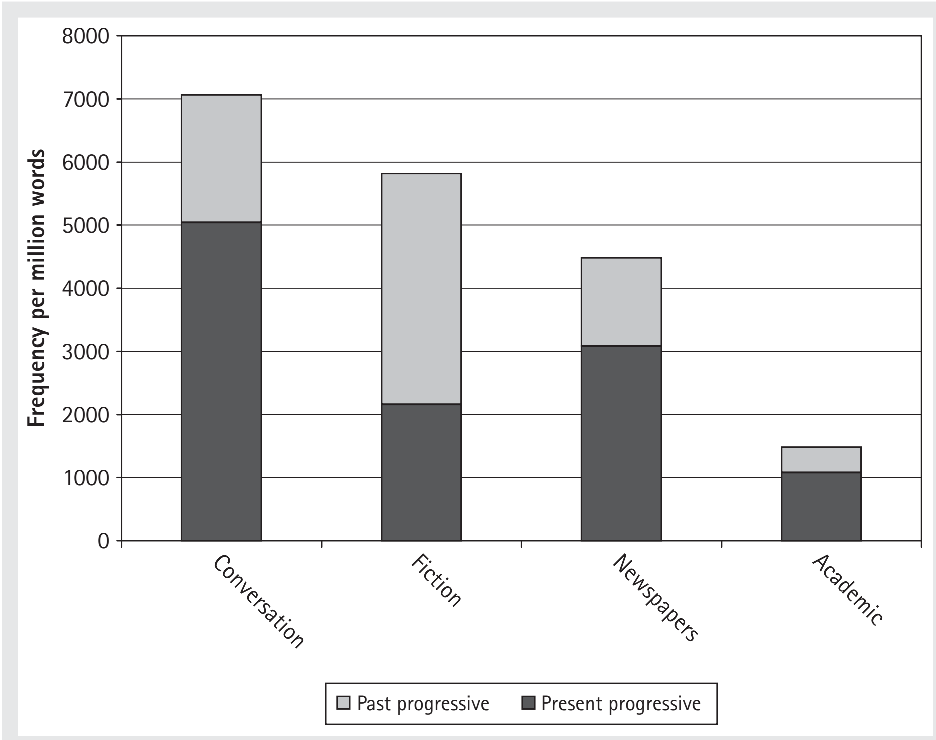


Figure 8.1. Distribution of progressive aspect verb phrases across registers

Kris Just do and be done with it. Smoke a joint <laugh>.
Jan Rita that'd help you sleep.
Rita No
Jan I don't think so.
Amanda They used to make me sleep.
Rita No that would make my mind race, yeah, typical.
Jan Okay let's do the Christmas.
Rita If I drink
Amanda Okay.
Rita If I smoke, anything, makes my mind race.
Amanda These tins are the last ones.
Jan It's just a little something Rita.
Rita You go overboard. Now, don't you make us feel guilty.

As the conversational excerpt above shows, verbs of all types tend to occur with simple aspect rather than progressive aspect, including stative relational verbs (e.g., *be*), mental verbs (e.g., *know, prefer, feel, think*), verbs of facilitation or causation

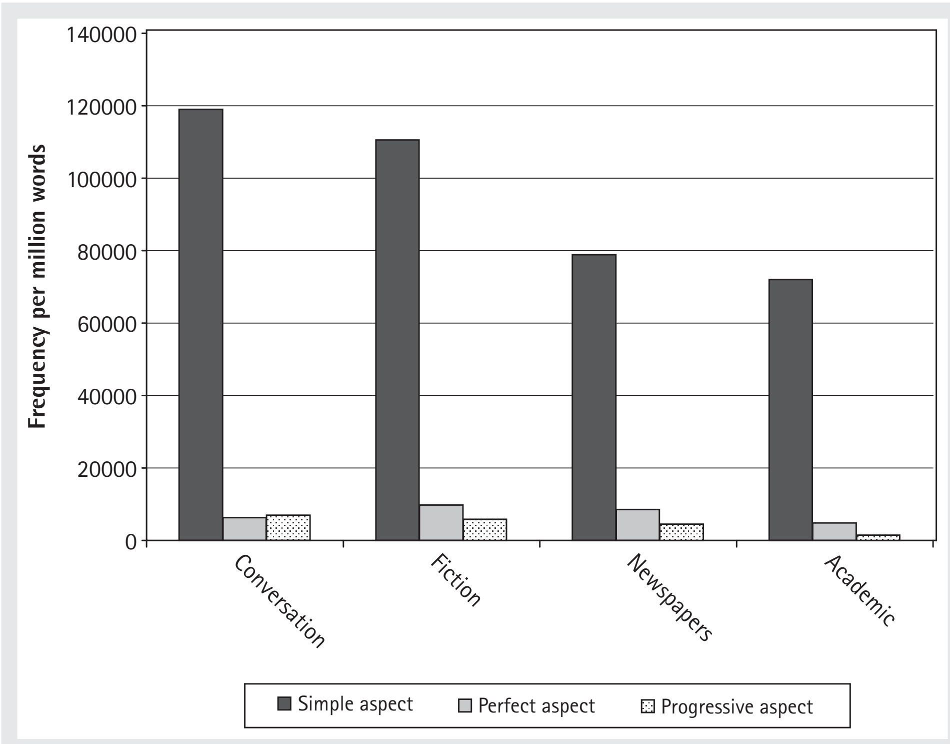


Figure 8.2. Distribution of aspect types across registers

(e.g., *let, help, make*), and activity verbs (e.g., *do, open, fall, roll, wait, smoke, sleep, race, drink, go*). There are a few particular verbs that occur more often with progressive aspect than simple aspect, such as *bleeding, chasing, shopping, dancing, dripping, marching, raining, sweating, chatting, joking, moaning, looking forward to, studying, lurking* (see Biber et al. 1999: 471–5). However, the normal style of discourse in conversation relies on simple aspect verbs (usually present tense), with shifts into progressive aspect being used to mark specialized meanings.

A second case study—focusing on dependent clause types—illustrates how corpus-based research has established the centrality of register for descriptions of language use. Dependent clauses are often regarded as one of the best measures of grammatical complexity. In some approaches, all dependent clause types are grouped together as manifesting complexity, as with the use of t-unit length to measure language development. Further, there is a strong expectation that writing manifests a much greater use of dependent clauses than speech. So, for example, students are expected to develop increasing use of dependent clauses as they progress in their academic writing skills (see, for example, Wolfe-Quintero et al. 1998).

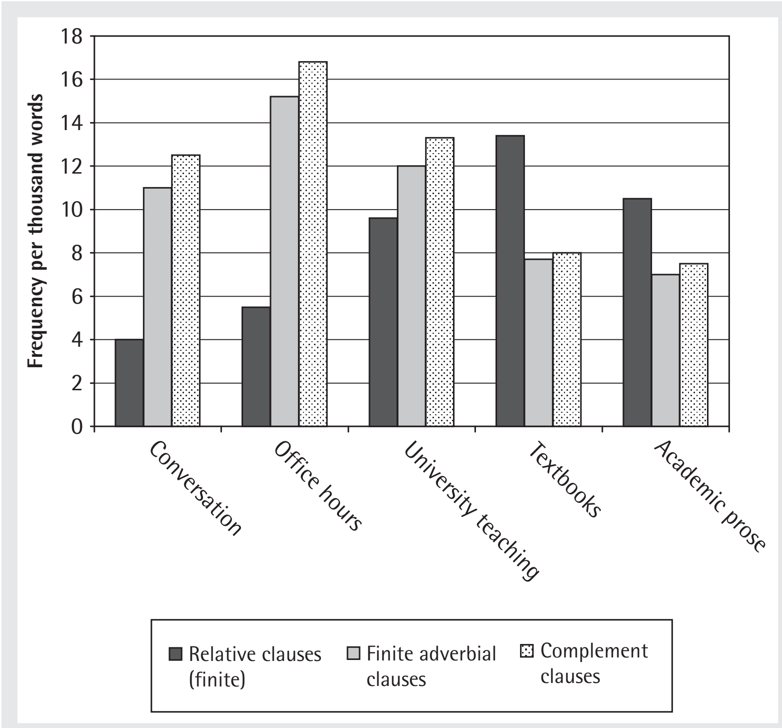


Figure 8.3. Distribution of dependent clause types across registers

Corpus-based research has shown that these predictions are based on faulty intuitions about use. That is, different dependent clause types are used and distributed in dramatically different ways, and some dependent clause types are actually much more common in conversation than in academic writing. Thus, the practice of treating all types of dependent clause as a single unified construct has no basis in actual language use.

For example, Figure 8.3 compares the use of dependent clause types in five spoken and written registers: conversation, university office hours, university teaching, university textbooks, and academic prose. Relative clauses follow the expected pattern of being much more common in academic writing and textbooks than in conversation (and office hours). Class teaching is intermediate between conversation and academic writing in the use of relative clauses. However, the other two clause types—adverbial clauses and complement clauses—are much more common in conversation than in academic writing. Office hours are interesting here because they are even more sharply distinguished from writing, with extremely frequent use of adverbial clauses and complement clauses. Class teaching is very similar to conversation in the frequent use of complement clauses and finite adverbial clauses.

Closer consideration of these patterns shows that they are interpretable in functional terms. For example, in conversation both adverbial and complement clauses occur with a highly restricted range of forms. Most adverbial clauses in conversation are finite, with especially high frequencies of *if*-clauses and *because*-clauses. Similarly, most complement clauses in conversation are finite (*that*-clauses and WH-clauses). In most cases, these complement clauses are controlled by a verb that expresses a "stance" relative to the proposition contained in the complement clause (e.g., *I thought that* ..., *I don't know why* ...).

In general, these distributional patterns conform to the general reliance on clausal rather than phrasal syntax in conversation (see Biber and Conrad to appear) and the communicative purposes of focusing on personal experience and activities rather than conveying more abstract information. These kinds of findings are typical of other corpus-based research, showing how the patterns of linguistic variation are systematically distributed in ways that have clear functional interpretations but are often not anticipated ahead of time.

8.3 Corpus-driven research studies

While corpus-based studies uncover surprising patterns of variation, corpus-driven analyses exploit the potential of a corpus to identify linguistic categories and units that have not been previously recognized. That is, in a corpus-driven analysis, the "descriptions aim to be comprehensive with respect to corpus evidence" (Tognini-Bonelli and Elena 2001: 84), so that even the "linguistic categories" are derived "systematically from the recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Bonelli and Elena 2001: 87).

In its most extreme form, the corpus-driven approach assumes only the existence of word forms; grammatical classes and syntactic structures have no a priori status in the analysis. In fact, even inflected variants of the same lemma are treated separately, with the underlying claim that each word form has its own grammar and its own meanings. So, for example, Stubbs (1993: 16) cites the example of *eye* vs. *eyes*, taken from Sinclair (1991*b*). The plural form *eyes* often refers to the physical body part and is modified by an attributive adjective (e.g., *blue eyes*) or a possessive determiner (e.g., *your eyes*). In contrast, the singular form rarely refers to a specific body part but is commonly used in fixed expressions, like *make eye contact, keep an eye on/out, catch your eye, in my mind's eye*. Thus, some corpus-driven research has challenged the utility of the notion of lemma, arguing instead that each word form tends to occur in distinctive grammatical contexts and tends to have distinct meanings and uses.

In actual practice, a fairly wide range of methodologies have been used under the umbrella of corpus-driven research. These methodologies can all be distinguished from corpus-based research by the nature of their central research goals:

- corpus-driven research: attempting to uncover new linguistic constructs through inductive analysis of corpora;
- corpus-based research: attempting to describe the systematic patterns of variation and use for linguistic features and constructs that have been previously identified by linguistic theory.

However, corpus-driven methodologies can differ from one study to the next in three key respects:

- the extent to which they are based on analysis of lemmas vs. each word form;
- the extent to which they are based on previously defined linguistic constructs (e.g., part-of-speech categories and syntactic structures) vs. simple sequences of words;
- the role of frequency evidence in the analysis.

The following sections survey some major corpus-driven studies, introducing the contributions that result from this research approach while also describing the key methodological differences within this general approach. Section 8.3.1 illustrates one specific type of analysis undertaken from an extreme corpus-driven approach: the investigation of "lexical bundles", which are the most common recurrent sequences of word forms in a register. It turns out that these word sequences have distinctive structural and functional correlates, even though they rarely correspond to complete linguistic structures recognized by current linguistic theories.

Next, section 8.3.2 surveys research done within the framework of "pattern grammar". These studies adopt a more hybrid approach: they assume the existence of some grammatical classes (e.g., verb, noun) and basic syntactic structures, but they are corpus-driven in that they focus on the linguistic units that emerge from corpus analysis, with a primary focus on the inter-relation of words, grammar, and meaning. Frequency plays a relatively minor role in analyses done within this framework. In fact, as discussed in section 8.3.3, there is somewhat of a disconnect between theoretical discussions of the corpus-driven approach, where analyses are based on "recurrent patterns" and "frequency distributions" (Tognini-Bonelli 2001: 87), and the actual practice of scholars working in pattern grammar, which has focused much more on form–meaning associations with relatively little accountability to quantitative evidence from the corpus.

Finally, section 8.3.4 introduces Multi-Dimensional analysis, which might also be considered a hybrid approach: it assumes the validity of predefined grammatical categories (e.g., nominalizations, past tense verbs) and syntactic features (e.g., WH relative clauses, conditional adverbial clauses), but it uses frequency-based corpus-driven methods to discover the underlying parameters of linguistic variation that best distinguish among spoken and written registers.

8.3.1 Lexical bundles

As noted above, the strictest form of corpus-driven analysis assumes only the existence of word forms. Some researchers interested in the study of formulaic language have adopted this approach, beginning with simple word forms and giving priority to frequency, to identify recurrent word sequences (e.g., Salem 1987; Altenberg and Eeg-Olofsson 1990; Altenberg 1998; Butler 1998; and Schmitt et al. 2004). Several of these studies have investigated recurrent word sequences under the rubric of "lexical bundles", comparing their characteristics in different spoken and written registers (e.g., Biber et al. 1999, Chapter 13; Biber and Conrad 1999; Biber et al. 2004; Cortes 2002; 2004; Partington and Morley 2004; Nesi and Basturkmen 2006; Biber and Barbieri 2007; Tracy-Ventura et al. 2007; and Biber et al. to appear).

Lexical bundles are defined as the multi-word sequences that recur most frequently and are distributed widely across different texts. Lexical bundles in English conversation are word sequences like *I don't know if* or *I just wanted to*. They are usually neither structurally complete nor idiomatic in meaning.

The initial analysis of lexical bundles in English (Biber et al. 1999, Chapter 13) compared the frequent word sequences in conversation and academic prose, based on analysis of *c*.5-million-word sub-corpora from each register. Figure 8.4 shows the overall distribution of all 3-word and 4-word lexical bundles occurring more than 10 times per million words (distributed across at least five different texts). Not surprisingly, there are almost 10 times as many 3-word bundles as 4-word bundles. It is perhaps more surprising that there are many more lexical bundles in conversation than in academic writing.

Lexical bundles are identified using a corpus-driven approach, based solely on distributional criteria (rate of occurrence of word sequences and their distribution across texts). As a result, lexical bundles are not necessarily complete structural units recognized by current linguistic theories. However, once they have been identified using corpus-driven techniques, it is possible to carry out an interpretive analysis to determine if they have any systematic structural and functional characteristics.

This *post-hoc* analysis shows that lexical bundles differ from the formulaic expressions identified using traditional methods in three major respects. First, lexical bundles are by definition extremely common. Second, most lexical bundles are not idiomatic in meaning and not perceptually salient. For example, the meanings of bundles like *do you want to* or *I don't know what* are transparent from the individual words. And, finally, lexical bundles usually do not represent a complete structural unit. For example, Biber et al. (1999: 993–1000) found that only 15% of the lexical bundles in conversation can be regarded as complete phrases or clauses, while less than 5% of the lexical bundles in academic prose represent complete structural units. Instead, most lexical bundles bridge two structural units: they begin at a clause or phrase boundary, but the last words of the bundle are

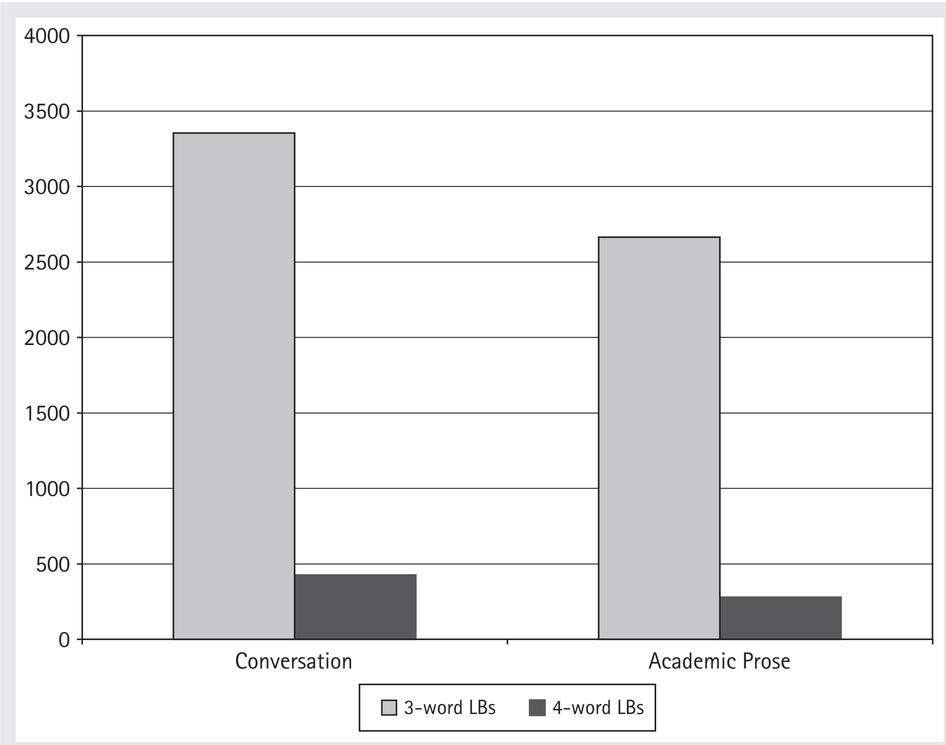


Figure 8.4. Number of different lexical bundles in English (occurring more than 10 times per million words)

the beginning elements of a second structural unit. Most of the bundles in speech bridge two clauses (e.g., *I want to know, well that's what I*), while bundles in writing usually bridge two phrases (e.g., *in the case of, the base of the*).

In contrast, the formulaic expressions recognized by linguistic theory are usually complete structural units and idiomatic in meaning. However, corpus analysis shows that formulaic expressions with those characteristics are usually quite rare. For example, idioms such as *kick the bucket* and *a slap in the face* are rarely attested in natural conversation. (Idioms are occasionally used in fictional dialogue, but even there they are not common; see Biber et al. 1999: 1024–6).

Although most lexical bundles are not complete structural units, they do usually have strong grammatical correlates. For example, bundles like *you want me to* are constructed from verbs and clause components, while bundles like *in the case of* are constructed from noun phrase and prepositional phrase components. In English, two major structural types of lexical bundle can be distinguished: clausal and phrasal. Many clausal bundles simply incorporate verb phrase fragments, such as *it's going to be* and *what do you think*. Other clausal bundles are composed of dependent clause fragments rather than simple verb phrase fragments, such as *when we get* *to* and *that I want to*. In contrast, phrasal bundles either consist of noun phrase components, usually ending with the start of a postmodifier (e.g., *the end of the, those of you who*), or prepositional phrase components with embedded modifiers (e.g., *of the things that*).

Figure 8.5 plots the distribution of these lexical bundle types across registers, showing that the structural correlates of lexical bundles in conversation are strikingly different from those in academic prose. (Figure 8.5 is based on a detailed analysis of the 4-word bundles that occur more than 40 times per million words.) In conversation, almost 90% of all common lexical bundles are declarative or interrogative clause segments. In fact, *c*.50% of these lexical bundles begin with a personal pronoun + verb phrase (such as *I don't know why*, *I thought that was*). An additional 19% of the bundles consist of an extended verb phrase fragment (e.g., *have a look at*), while another 17% of the bundles are question fragments (e.g., *can I have a*). In contrast, the lexical bundles in academic prose are phrasal rather than clausal. Almost 70% of the common bundles in academic prose consist of a noun phrase with an embedded prepositional phrase fragment (e.g., *the nature of the*) or a sequence that bridges across two prepositional phrases (e.g., *as a result of*).

Although they are neither idiomatic nor structurally complete, lexical bundles are important building blocks in discourse. Lexical bundles often provide a kind of pragmatic "head" for larger phrases and clauses; the bundle functions as a discourse frame for the expression of new information in the following slot. That is, the lexical bundle usually expresses stance or textual meanings, while the remainder of the phrase/clause expresses new propositional information that has been framed by the lexical bundle. In this way, lexical bundles provide interpretive frames for the developing discourse. For example,

> *I want you to write a very brief summary of his lecture.*
> *Hermeneutic efforts are provoked by the fact that the interweaving of system integration and social integration [. . .] keeps societal processes transparent . . .*

Three primary discourse functions can be distinguished for lexical bundles in English: (1) stance expressions, (2) discourse organizers, and (3) referential expressions (see Biber et al. 2004). Stance bundles express epistemic evaluations or attitudinal/modality meanings:

> ***Epistemic lexical bundles***:
> *I don't know what the voltage is here.*
> *I thought it was the other way around.*
>
> ***Attitudinal/modality bundles***:
> *I don't want to deliver bad news to her.*
> *All you have to do is work on it.*

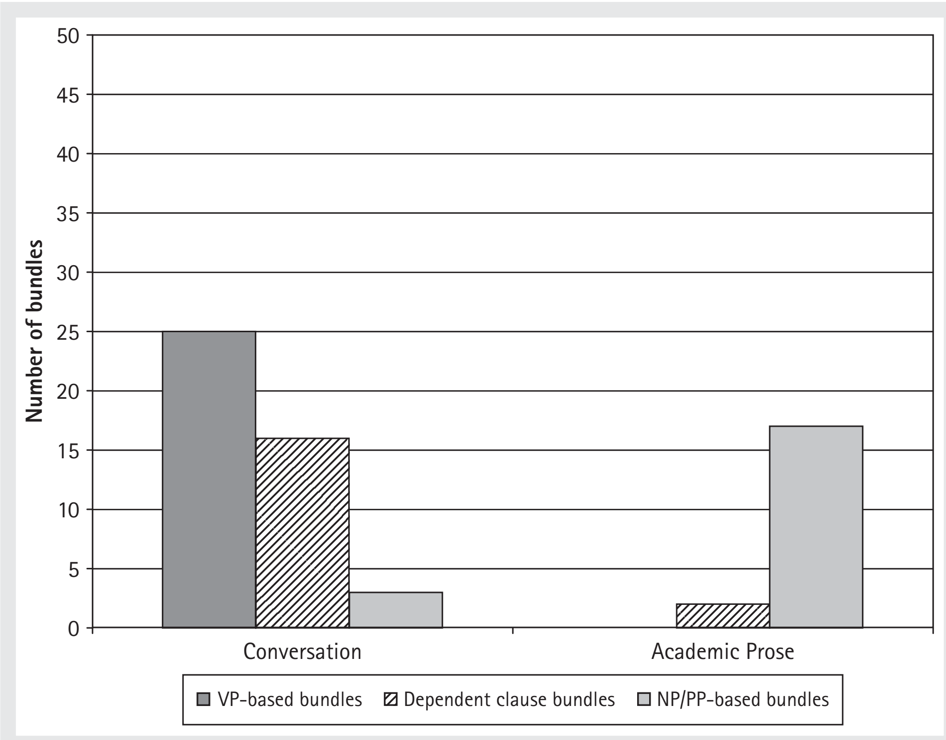


Figure 8.5. Distribution of lexical bundles across structural types (4-word bundles occurring more than 40 times per million words)

Discourse-organizing bundles function to indicate the overall discourse structure: introducing topics, topic elaboration/clarification, confirmation checks, etc.:

What I want to do is quickly run through the exercise . . .
Yes, you know there was more of a playful thing with it, you know what I mean?

Finally, referential bundles specify an entity or single out some particular attribute of an entity as especially important:

Students must define and constantly refine the nature of the problem.
She's in that office down there, at the end of the hall.

Figure 8.6 shows that the typical discourse functions of lexical bundles are strikingly different in conversation vs. academic writing: most bundles are used for stance functions in conversation, with a number also being used for discourse-organizing functions. In contrast, most bundles are used for referential functions in academic prose. These findings indicate that formulaic expressions develop to serve the most important communicative needs of a register. It further turns out

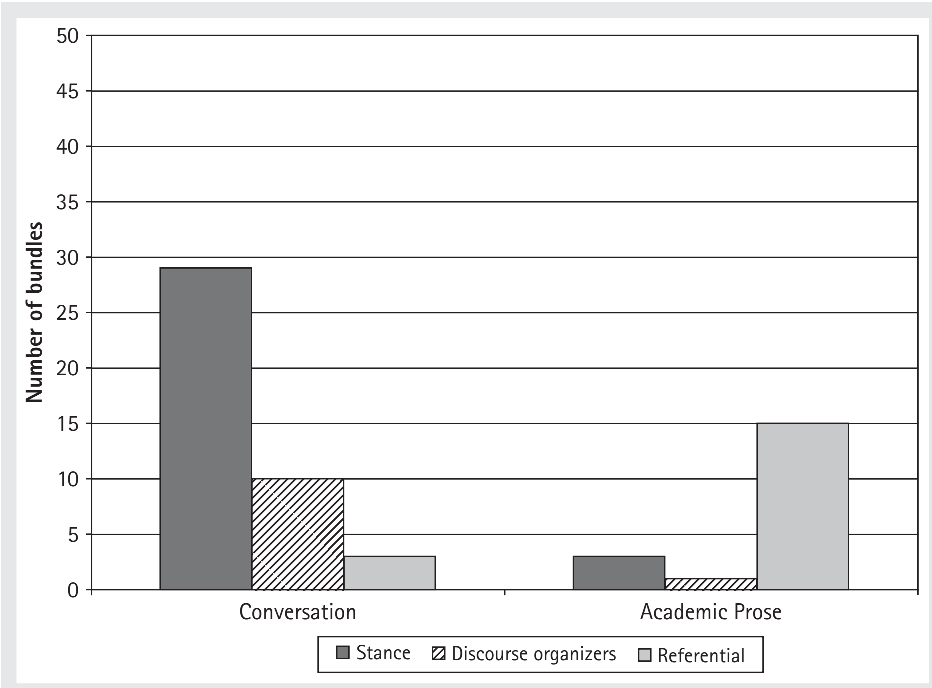


Figure 8.6. Distribution of lexical bundles across functional types (4-word bundles occurring more than 40 times per million words)

that there is a strong association between structural type and functional type for these lexical bundles: most stance bundles employ verbs or clause fragments, while most referential bundles are composed of noun phrase and prepositional phrase fragments.

In summary, a minimalist corpus-driven approach, beginning with only the existence of word forms, shows that words in English co-occur in highly frequent fixed sequences. These sequences are not complete constituents recognized by traditional theories, but they are readily interpretable in both structural and functional terms.

8.3.2 The interdependence of lexis, grammar, and meaning: Pattern grammar

Many scholars working within a corpus-driven framework have focused on the meaning and use of particular words, arguing that lexis, grammar, and meaning are fundamentally intertwined (e.g., Francis et al. 1996; 1998; Hunston and Francis 1998; 2000; Sinclair 1991*a*; Stubbs 1993; and Tognini-Bonelli 2001). The best-developed application of corpus-driven research with these goals is the "pattern grammar" reference book series (e.g., Francis et al. 1996; 1998; see also Hunston and Francis 2000).

The pattern grammar studies might actually be considered hybrids, combining corpus-based and corpus-driven methodologies. They are corpus-based in that they assume the existence (and definition) of basic part-of-speech categories and some syntactic constructions, but they are corpus-driven in that they focus primarily on the construct of the grammatical *pattern*: "a phraseology frequently associated with (a sense of) a word... Patterns and lexis are mutually dependent, in that each pattern occurs with a restricted set of lexical items, and each lexical item occurs with a restricted set of patterns. In addition, patterns are closely associated with meaning, firstly because in many cases different senses of words are distinguished by their typical occurrence in different patterns; and secondly because words which share a given pattern tend also to share an aspect of meaning" (Hunston and Francis 2000: 3). Thus, a pattern is a combination of words that "occurs relatively frequently", is "dependent on a particular word choice", and has "a clear meaning associated with it" (Hunston and Francis 2000: 37). Grammatical patterns are *not* necessarily complete structures (phrases or clauses) recognized by linguistic theory. Thus, following the central defining characteristic of corpus-driven research given above, the pattern grammar studies attempt to uncover new linguistic constructs—the *patterns*—through inductive analysis of corpora.

A central claim of this framework is that grammatical patterns have inherent meaning, shared across the set of words that can occur in a pattern. For example, many of the verbs that occur in the grammatical pattern V+ *over* +NP express meanings relating to conflict or disagreement, such as *bicker, disagree, fight, quarrel, quibble*, and *wrangle* (see Hunston and Francis 2000: 43–4); thus it can be argued that the grammatical pattern itself somehow entails this meaning.

The pattern grammar reference books (Francis et al. 1996; 1998) have attempted to provide a comprehensive catalog of the grammatical patterns for verbs, nouns, and adjectives in English. These books show that there are systematic regularities in the associations between grammatical frames, sets of words, and particular meanings on a much larger scale than it could have been possible to anticipate before the introduction of large-scale corpus analysis. For example, the reference book on grammatical patterns for verbs (Francis et al. 1996) includes over 700 different patterns and catalogs the use of over 4,000 verbs with respect to those patterns. The reference book on grammatical patterns for nouns and adjectives (Francis et al. 1998) is similar in scope, with over 200 patterns used to describe the use of over 8,000 nouns and adjectives.

The pattern grammar reference books do not address some of the stronger theoretical claims that have been associated with the corpus-driven approach. For example, "patterns" are based on analysis of lemmas rather than individual word forms, and thus the pattern grammar studies provide no support for the general claim that each word form has its own grammar.[1]

The pattern grammar studies also do not support the strong version of the claim that each grammatical pattern has its own meaning. In fact, it is rarely the case that a grammatical frame corresponds to a single meaning domain. However, these studies do provide extensive support for a weaker form of the claim, documenting how the words that occur in a grammatical frame belong to a relatively small set of meaning groups. For example, the adjectives that occur in the grammatical frame **ADJ *in* N** mostly fall into several major meaning groups, such as:

- adjectives that express high interest or participation:
  e.g., *absorbed, embroiled, engaged, engrossed, enmeshed, immersed, interested, involved, mixed up, wrapped up*
- adjectives that express a deficit:
  e.g., *deficient, lacking, wanting*
- adjectives that express an amount or degree:
  e.g., *awash, high, low, poor, rich*
- adjectives that express proficiency or fluency
  e.g., *fluent, proficient, schooled, skilful, skilled, versed*
- adjectives that express that something is covered
  e.g., *bathed, clad, clothed, coated, plastered, shrouded, smothered*
  (see Francis et al. 1998: 444–51; Hunston and Francis 2000: 75–6).

As noted above, the methodology used for the pattern grammar studies relaxes the strict requirements of corpus-driven methodology. First, pre-defined grammatical constructs are used in the approach, including basic grammatical classes, phrase types, and even distinctions that require a priori syntactic analysis. In addition, frequency plays only a minor role in the analysis, and some word combinations that occur frequently are not regarded as patterns at all. For example, the nouns followed by complementizer *that* are analyzed as patterns (e.g., *fact, claim, stipulation, expectation, disgust, problem*, etc.), but nouns followed by the relative pronoun *that* do not constitute a pattern, even if the combination is frequent (e.g., *extent, way, thing, questions, evidence, factors + that*). Similarly, prepositions are analyzed for their syntactic function in the sequence noun + preposition, to distinguish between prepositional phrases functioning as adverbials (which do not count as part of any pattern), vs. prepositional phrases that complement the preceding noun

[1] Other studies that advocate this position have been based on a few selected case studies (e.g., Sinclair 1991*b* on *eye* vs. *eyes*; Tognini-Bonelli and Elena 2001: 92–8 on *facing* vs. *faced*, and *saper* vs. *sapere* in Italian). These case studies clearly show that word forms belonging to the same lemma do sometimes have their own distinct grammar and meaning. However, no empirical study to date has investigated the extent to which this situation holds across the full set of word forms and lemmas in a language. (In contrast, the pattern grammar reference books seem to implicitly suggest that most inflected word forms that belong to a single lemma "pattern" in similar ways.)

(which do constitute a pattern). So, for example, the combinations for the pattern **ADJ *in* N** listed above all include a prepositional phrase that complements the adjective. In contrast, when the prepositional phrase has an adverbial function, it is analyzed as *not* representing a pattern, even if the combination is frequent. Thus, the following adjectives do not belong to any pattern when they occur in the combination **ADJ *in* N**, even though they occur frequently and represent relatively coherent meaning groups:

adamant, firm, resolute, steadfast, unequivocal
loud, vehement, vocal, vociferous
(see Hunston and Francis 2000: 76).

Regardless of the specific methodological considerations, the corpus-driven approach as realized in the pattern grammar studies has shown that there are systematic regularities in the associations between grammatical frames, sets of words, and particular meanings, on a much more comprehensive scale than it could have been possible to anticipate before the availability of large corpora and corpus-analysis tools.

8.3.3 The role of frequency in corpus-driven analysis

Surprisingly, one major difference among corpus-driven studies concerns the role of frequency evidence. Nearly every description of the corpus-driven approach includes mention of frequency, as in: (a) the "linguistic categories" are derived "systematically from the recurrent patterns and the frequency distributions that emerge from language in context" (Tognini-Bonelli 2001: 87); (b) in a grammar pattern, "a combination of words occurs relatively frequently" (Hunston and Francis 2000: 37).

In the study of lexical bundles, frequency evidence is primary. This framework can be regarded as the most extreme test of the corpus-driven approach, addressing the question of whether the most commonly occurring sequences of word forms can be interpreted as linguistically significant units. In contrast, frequency is not actually important in pattern grammar studies. On the one hand, frequent word combinations are not included in the pattern analysis if they represent different syntactic constructions, as described in the last section. The combination *satisfaction that* provides another example of this type. When the *that* initiates a complement clause, this combination is one of the realizations of the "happiness" **N *that*** pattern (Francis et al. 1998: 111), as in:

One should of course record one's <u>satisfaction that</u> the two leaders got on well together.

However, it is much more frequent for the combination *satisfaction that* to represent different syntactic constructions, as in:

(a) *The satisfaction provided by conformity is in competition with the often more immediate satisfaction that can be provided by crime.*
(b) *He then proved to his own satisfaction that all such endeavours were doomed to failure.*

In (a), the word *that* initiates a relative clause, and in (b), the *that* initiates a verb complement clause controlled by *proved*. Neither of these combinations are analyzed as belonging to a pattern, even though they are more frequent than the combination of *satisfaction* followed by a *that* noun complement clause.

Thus, frequency is not a decisive factor in identifying "patterns", despite the definition that requires that the combination of words in a pattern must occur "relatively frequently". Instead, the criteria that a grammatical pattern must be associated with a particular set of words and have a clear meaning are more decisive (see Hunston and Francis 2000: 67–76).

In fact, some corpus-driven linguists interested in the lexis–grammar interface have overtly argued against the importance of frequency. For example, Sinclair notes that

> some numbers are more important than others. Certainly the distinction between 0 and 1 is fundamental, being the occurrence or non-occurrence of a phenomenon. The distinction between 1 and more than one is also of great importance . . . [because even two unconnected tokens constitute] the recurrence of a linguistic event . . . , [which] permits the reasonable assumption that the event can be systematically related to a unit of meaning. In the study of meaning it is not usually necessary to go much beyond the recognition of recurrence [i.e., two independent tokens]. . . . (Sinclair 2001: 343–4)

Similarly, Tognini-Bonelli notes that

> It is therefore appropriate to set up as the minimum sufficient condition for a pattern of occurrence to merit a place in the description of the language, that it occurs at least twice, and the occurrences appear to be independent of each other . . . .
>
> (Tognini-Bonelli 2001: 89)

Thus, there is some tension here between the underlying definition of the corpus-driven approach, which derives linguistic categories from "recurrent patterns" and "frequency distributions" (Tognini-Bonelli 2001: 87), and the actual practice of scholars working on pattern grammar and the lexis–grammar–meaning interconnection, which has focused much more on form–meaning associations with relatively little accountability to quantitative distributional patterns in a corpus. Here again, we see the central defining characteristic of corpus-driven research to be the shared goal of identifying new linguistic constructs through inductive analysis of a corpus, regardless of differences in the specific methodological approaches.

8.3.4 Linguistic "dimensions" of register variation

As discussed in section 8.2 above, corpus research has been used to describe particular linguistic features and their variants, showing how these features vary in their distribution and patterns of use across registers. This relationship can also be approached from the opposite perspective, with a focus on describing the registers rather than describing the use of particular linguistic features.

It turns out, though, that the distribution of individual linguistic features cannot reliably distinguish among registers. There are simply too many different linguistic characteristics to consider, and individual features often have idiosyncratic distributions. Instead, sociolinguistic research has argued that register descriptions must be based on linguistic co-occurrence patterns (see, for example, Ervin-Tripp 1972; Hymes 1974; Brown and Fraser 1979: 38–9; Halliday 1988: 162).

Multi-Dimensional (MD) analysis is a corpus-driven methodological approach that identifies the frequent linguistic co-occurrence patterns in a language, relying on inductive empirical/quantitative analysis (see, for example, Biber 1988; 1995). Frequency plays a central role in the analysis, since each dimension represents a constellation of linguistic features that frequently co-occur in texts. These "dimensions" of variation can be regarded as linguistic constructs not previously recognized by linguistic theory. Thus, although the framework was developed to describe patterns of register variation (rather than the meaning and use of individual words), MD analysis is clearly a corpus-driven methodology in that the linguistic constructs—the "dimensions"—emerge from analysis of linguistic co-occurrence patterns in the corpus.

The set of co-occurring linguistic features that comprise each dimension is identified quantitatively. That is, based on the actual distributions of linguistic features in a large corpus of texts, statistical techniques (specifically factor analysis) are used to identify the sets of linguistic features that frequently co-occur in texts.

The original MD analyses investigated the relations among general spoken and written registers in English, based on analysis of the LOB (Lancaster–Oslo–Bergen) Corpus (15 written registers) and the London–Lund Corpus (six spoken registers). Sixty-seven different linguistic features were analyzed computationally in each text of the corpus. Then, the co-occurrence patterns among those linguistic features were analyzed using factor analysis, identifying the underlying parameters of variation: the factors or "dimensions". In the 1988 MD analysis, the 67 linguistic features were reduced to seven underlying dimensions. (The technical details of the factor analysis are given in Biber 1988, Chapters 4–5; see also Biber 1995, Chapter 5).

The dimensions are interpreted functionally, based on the assumption that linguistic co-occurrence reflects underlying communicative functions. That is, linguistic features occur together in texts because they serve related communicative functions.

The most important features on Dimensions 1–5 in the 1988 MD analysis are:

Dimension 1: Involved vs. Informational Production

Positive features: mental (private) verbs, *that* complementizer deletion, contractions, present tense verbs, WH-questions, 1st and 2nd person pronouns, pronoun *it*, indefinite pronouns, *do* as pro-verb, demonstrative pronouns, emphatics, hedges, amplifiers, discourse particles, causative subordination, sentence relatives, WH-clauses

Negative features: nouns, long words, prepositions, type/token ratio, attributive adjectives

Dimension 2: Narrative vs. Non-narrative Discourse

Positive features: past tense verbs, 3rd person pronouns, perfect aspect verbs, communication verbs

Negative features: present tense verbs, attributive adjectives

Dimension 3: Situation-dependent vs. Elaborated Reference

Positive features: time adverbials, place adverbials, other adverbs

Negative features: WH-relative clauses (subject gaps, object gaps), phrasal coordination, nominalizations

Dimension 4: Overt Expression of Argumentation

Positive features: prediction modals, necessity modals, possibility modals, suasive verbs, conditional subordination, split auxiliaries

Dimension 5: Abstract/Impersonal Style

Positive features: conjuncts, agentless passives, BY-passives, past participial adverbial clauses, past participial postnominal clauses, other adverbial subordinators

Each dimension can have "positive" and "negative" features. Rather than reflecting importance, positive and negative signs identify two groupings of features that occur in a complementary pattern as part of the same dimension. That is, when the positive features occur together frequently in a text, the negative features are markedly less frequent in that text, and vice versa.

On Dimension 1, the interpretation of the negative features is relatively straightforward. Nouns, word length, prepositional phrases, type/token ratio, and attributive adjectives all reflect an informational focus, a careful integration of information in a text, and precise lexical choice. Text Sample 1 illustrates these co-occurring linguistic characteristics in an academic article:

Text Sample 1. Technical academic prose

Apart from these very general group-related aspects, there are also individual aspects that need to be considered. Empirical data show that similar processes can be guided quite differently by users with different views on the purpose of the communication.

This text sample is typical of written expository prose in its dense integration of information: frequent nouns and long words, with most nouns being modified by attributive adjectives or prepositional phrases (e.g., *general group-related aspects, individual aspects, empirical data, similar processes, users with different views on the purpose of the communication*).

The set of positive features on Dimension 1 is more complex, although all of these features have been associated with interpersonal interaction, a focus on personal stance, and real-time production circumstances. For example, first and second person pronouns, WH-questions, emphatics, amplifiers, and sentence relatives can all be interpreted as reflecting interpersonal interaction and the involved expression of personal stance (feelings and attitudes). Other positive features are associated with the constraints of real time production, resulting in a reduced surface form, a generalized or uncertain presentation of information, and a generally "fragmented" production of text; these include *that*-deletions, contractions, pro-verb DO, the pronominal forms, and final (stranded) prepositions. Text Sample 2 illustrates the use of positive Dimension 1 features in a workplace conversation:

Text Sample 2. Conversation at a reception at work

SABRINA I'm dying of thirst.
SUZANNA Mm, hmm. Do you need some M & Ms?
SABRINA Desperately. <laugh> Ooh, thank you. Ooh, you're so generous.
SUZANNA Hey I try.
SABRINA Let me have my Snapple first. Is that cold–cold ?
SUZANNA I don't know but there should be ice on uh, <unclear>.
SABRINA I don't want to seem like I don't want to work and I don't want to seem like a stuffed shirt or whatever but I think this is really boring.
SUZANNA I know.
SABRINA I would like to leave here as early as possible today, go to our rooms, and pick up this thing at eight o'clock in the morning.
SUZANNA Mm, hmm.

Overall, Factor 1 represents a dimension marking interactional, stance-focused, and generalized content (the positive features mentioned earlier) vs. high informational density and precise word choice (the negative features). Two separate communicative parameters seem to be represented here: the primary purpose of the writer/speaker (involved vs. informational), and the production circumstances (those restricted by real-time constraints vs. those enabling careful editing possibilities). Reflecting both of these parameters, the interpretive label "Involved vs. Informational Production" was proposed for the dimension underlying this factor.

The second major step in interpreting a dimension is to consider the similarities and differences among registers with respect to the set of co-occurring linguistic features. To achieve this, *dimension scores* are computed for each text, by summing the individual scores of the features that co-occur on a dimension (see Biber 1988: 93–7). For example, the Dimension 1 score for each text was computed by adding together the frequencies of private verbs, *that*-deletions, contractions, present tense verbs, etc.—the features with positive loadings—and then subtracting the frequencies of nouns, word length, prepositions, etc.—the features with negative loadings.

Once a dimension score is computed for each text, the mean dimension score for each register can be computed. Plots of these mean dimension scores allow linguistic characterization of any given register, comparison of the relations between any two registers, and a fuller functional interpretation of the underlying dimension.

For example, Figure 8.7 plots the mean dimension scores of registers along Dimension 1 from the 1988 MD analysis. The registers with large positive values (such as face-to-face and telephone conversations), have high frequencies of present tense verbs, private verbs, first and second person pronouns, contractions, etc.—the features with salient positive weights on Dimension 1. At the same time, registers with large positive values have markedly low frequencies of nouns, prepositional phrases, long words, etc.—the features with salient negative weights on Dimension 1. Registers with large negative values (such as academic prose, press reportage and official documents) have the opposite linguistic characteristics: very high frequencies of nouns, prepositional phrases, etc., plus low frequencies of private verbs, contractions, etc.

The relations among registers shown in Figure 8.7 confirm the interpretation of Dimension 1 as distinguishing among texts along a continuum of involved vs. informational production. At the positive extreme, conversations are highly interactive and involved, with the language produced under real-time circumstances. Registers such as public conversations (interviews and panel discussions) are intermediate: they have a relatively informational purpose, but participants interact with one another and are still constrained by real time production. Finally, at the negative extreme, registers such as academic prose are non-interactive but highly informational in purpose, produced under controlled circumstances that permit extensive revision and editing.

Figure 8.7 shows that there is a large range of variation among spoken registers with respect to the linguistic features that comprise Dimension 1 ("Involved vs. Informational Production"). Conversation has extremely large positive Dimension 1 scores; spontaneous speeches and interviews have moderately large positive scores;

```
INVOLVED
      | TELEPHONE CONVERSATIONS
      |
   35 + FACE-TO-FACE CONVERSATIONS
      |
      |
      |
   30 +
      |
      |
      |
   25 +
      |
      |
      |
   20 + Personal letters
      | PUBLIC CONVERSATIONS, SPONTANEOUS SPEECHES
      | INTERVIEWS
      |
   15 +
      |
      |
      |
   10 +
      |
      |
      |
    5 +
      | Romance fiction
      | PREPARED SPEECHES
      |
    0 + Mystery and adventure fiction
      | General fiction
      | Professional letters
      | BROADCASTS
   –5 +
      | Science fiction
      | Religion
      | Humor
  –10 + Popular lore, editorials, hobbies
      |
      | Biographies
      | Press reviews
  –15 + Academic prose, press reportage
      |
      | Official documents
INFORMATIONAL
```

Figure 8.7. Mean scores of registers along Dimension 1: Involved vs. Informational Production (adapted from Figure 7.1 in Biber 1988)

Note: Underlining denotes written registers; capitalization denotes spoken registers; $F = 111.9$, $p < .0001$, $r^2 = 84.3\%$.

while prepared speeches and broadcasts have scores around 0.0 (reflecting a balance of positive and negative linguistic features on this dimension). The written registers similarly show an extensive range of variation along Dimension 1. Expository informational registers, like official documents and academic prose, have very large negative scores; the fiction registers have scores around 0.0; while personal letters have a relatively large positive score.

This distribution shows that no single register can be taken as representative of the spoken or written mode. At the extremes, written informational prose is dramatically different from spoken conversation with respect to Dimension 1 scores. But written personal letters are relatively similar to spoken conversation, while spoken prepared speeches share some Dimension 1 characteristics with written fictional registers. Taken together, these Dimension 1 patterns indicate that there is extensive overlap between the spoken and written modes in these linguistic characteristics, while the extremes of each mode (i.e., conversation vs. informational prose) are sharply distinguished from one another.

The overall comparison of speech and writing resulting from the 1988 MD analysis is actually much more complex because six separate dimensions of variation were identified and each of these defines a different set of relations among spoken and written registers. For example, Dimension 2 is interpreted as "Narrative vs. Non-narrative Concerns". The positive features—past tense verbs, third person pronouns, perfect aspect verbs, communication verbs, and present participial clauses—are associated with past time narration. In contrast, the positive features—present tense verbs and attributive adjectives—have non-narrative communicative functions.

The distribution of registers along Dimension 2, shown in Figure 8.8, further supports its interpretation as Narrative vs. Non-narrative Concerns. All types of fiction have markedly high positive scores, reflecting their emphasis on narrating events. In contrast, registers which are typically more concerned with events currently in progress (e.g., broadcasts) or with building arguments rather than narrating (e.g., academic prose) have negative scores on this dimension. Finally, some registers have scores around 0.0, reflecting a mix of narrative and other features. For example, face-to-face conversation will often switch back and forth between narration of past events and discussion of current interactions.

Each of the dimensions in the analysis can be interpreted in a similar way. Overall, the 1988 MD analysis showed that English registers vary along several underlying dimensions associated with different functional considerations, including: interactiveness, involvement and personal stance, production circumstances, informational density, informational elaboration, narrative purposes, situated reference, persuasiveness or argumentation, and impersonal presentation of information.

```
NARRATIVE
        | Romance fiction
     7  +
        |
        |
        |
     6  +  Mystery, science, and general fiction
        |
        |  Adventure fiction
        |
     5  +
        |
        |
        |
     4  +
        |
        |
        |
     3  +
        |
        |
        |
     2  + Biographies
        |
        |  SPONTANEOUS SPEECHES
        |
     1  + Humor
        | PREPARED SPEECHES
        | Press reportage
        | Personal letters
     0  + Popular lore
        |
        | FACE-TO-FACE CONVERSATIONS
        | Religion; Editorials
    –1  + PUBLIC CONVERSATIONS
        |
        | Press reviews
        |
    –2  + TELEPHONE CONVERSATIONS
        | Professional letters
        | Academic prose
        | Official documents
    –3  + Hobbies
        | BROADCASTS
        |
NON-NARRATIVE
```

Figure 8.8. Mean scores for registers along Dimension 2: Narrative vs. Non-Narrative Discourse (adapted from Figure 7.2 in Biber 1988)

Note: Underlining denotes written registers; capitalization denotes spoken registers; $F = 32.3$, $p < .0001$, $r^2 = 60.8\%$.

Many studies have applied the 1988 dimensions of variation to study the linguistic characteristics of more specialized registers and discourse domains. For example:

Present-day registers:	*Studies*:
spoken and written university registers	Biber et al. (2002)
AmE vs. BrE written registers	Biber (1987)
AmE vs. BrE conversational registers	Helt (2001)
biology vs. history student and academic writing	Conrad (1996; 2001)
I-M-R-D sections in medical research articles	Biber and Finegan (1994*b*)
direct mail letters	Connor and Upton (2003)
discourse moves in non-profit grant proposals	Connor and Upton (2004)
oral proficiency interviews	Connor-Linton and Shohamy (2001)
academic lectures	Csomay (2005)
conversation vs. TV dialogue	Quaglio (2009)
female/male conversational style	Rey (2001); Biber and Burges (2000)
author styles	Connor-Linton (2001); Biber and Finegan (1994*a*)

Historical registers:	*Studies*:
written and speech-based registers; 1650–present	Biber and Finegan (1989; 1997)
medical research articles and scientific research articles; 1650–present	Atkinson (1992; 1996; 1999)
19th-century written registers	Geisler (2002)

However, other MD studies have undertaken new corpus-driven analyses to identify the distinctive sets of co-occurring linguistic features that occur in a particular discourse domain or in a language other than English. The following section surveys some of those studies.

8.3.4.1 *Comparison of the multi-dimensional patterns across discourse domains and languages*

Numerous other studies have undertaken complete MD analyses, using factor analysis to identify the dimensions of variation operating in a particular discourse domain in English, rather than applying the dimensions from the 1988 MD analysis (e.g., Biber 1992; 2001; 2006; 2008; Biber and Jones 2006; Biber et al. 2007; Friginal 2008; 2009; Kanoksilapatham 2007; Crossley and Louwerse 2007; Reppen 2001).

Given that each of these studies is based on a different corpus of texts, representing a different discourse domain, it is reasonable to expect that they would each identify a unique set of dimensions. This expectation is reinforced by the fact that the more recent studies have included additional linguistic features not used in earlier MD studies (e.g., semantic classes of nouns and verbs). However, despite these differences in design and research focus, there are certain striking similarities in the set of dimensions identified by these studies.

Most importantly, in nearly all of these studies, the first dimension identified by the factor analysis is associated with an informational focus vs. a personal focus (personal involvement/stance, interactivity, and/or real-time production features). For example:

Study and dimension	*Corpus*	*Linguistic features defining the dimension*
Biber (2001) Dimension 1	18th-c. written and speech-based registers	prepositions, passives, nouns, long words, past tense verbs vs. 1st and 2nd person pronouns, present tense, possibility and prediction modals, *that*-deletion, mental verbs, emphatics
Biber (2006) Dimension 1	university spoken and written registers	nominalizations, long words, nouns, prepositions, abstract nouns, attributive adjectives, passives, stance noun + *to*-clause, etc. vs. contractions, demonstrative pronouns, *it*, 1st person pronouns, present tense, time advs, *that*-omission, WH-questions, etc.
White (1994) Dimension 1	job interviews	long words, nouns, nominalizations, prepositions, WH-questions, 2nd person pronouns vs. 1st person pronouns, contractions, adverbs, discourse particles, emphatics, etc.
Reppen (2001) Dimension 1	elementary school registers	nouns, long words, nominalizations, passives, attributive adjectives, prepositions vs. initial *and*, time adverbials, 3rd person pronouns
Biber (2008) Dimension 1	conversational text types	long words, nominalizations, prepositions, abstract nouns, relative clauses, attributive adjs. vs. contractions, 1st and 2nd person pronouns, activity verbs

It is perhaps not surprising that Dimension 1 in the original 1988 MD analysis was strongly associated with an informational vs. (inter)personal focus, given that the corpus in that study ranged from spoken conversational texts to written expository texts. For the same reason, it is somewhat predictable that a similar dimension would have emerged from the study of 18th-century written and speech-based registers. It is somewhat more surprising that academic spoken and written registers would be defined by a similar linguistic dimension (and especially surprising that classroom teaching is similar to conversation, and strikingly different from academic writing, in the use of these linguistic features). And it was completely unexpected that a similar oral/literate dimension—realized by essentially the same set of co-occurring linguistic features—would be fundamentally important in highly restricted discourse domains, including studies of job interviews, elementary school registers, and variations among the different kinds of conversation.

A second parameter found in most MD analyses corresponds to narrative discourse, reflected by the co-occurrence of features like past tense, third person pronouns, perfect aspect, and communication verbs (see, for example, the Biber 2006 study of university registers; Biber 2001 on 18th-century registers; and the Biber 2008 study of conversation text types). In some studies, a similar narrative dimension emerged with additional special characteristics. For example, in Reppen's (2001) study of elementary school registers, "narrative" features like past tense, perfect aspect, and communication verbs co-occurred with once-occurring words and a high type/token ratio; in this corpus, history textbooks rely on a specialized and diverse vocabulary to narrate past events. In the job interview corpus (White 1994), the narrative dimension reflected a fundamental opposition between personal/specific past events and experiences (past tense verbs co-occurring with first person singular pronouns) vs. general practice and expectations (present tense verbs co-occurring with first person plural pronouns). In Biber and Kurjian's (2007) study of web text types, narrative features co-occurred with features of stance and personal involvement on the first dimension, distinguishing personal narrative web pages (e.g., personal blogs) from the various kinds of more informational web pages.

At the same time, most of these studies have identified some dimensions that are unique to the particular discourse domain. For example, the factor analysis in Reppen (1994) identified a dimension of "Other-directed idea justification" in elementary student registers. The features on this dimension include second person pronouns, conditional clauses, and prediction modals; these features commonly co-occur in certain kinds of student writings (e.g., *If you wanted to watch TV a lot you would not get very much done*).

The factor analysis in Biber's (2006) study of university spoken and written registers identified four dimensions. Two of these are similar linguistically and functionally to dimensions found in other MD studies: Dimension 1: "Oral vs. literate discourse"; and Dimension 3: "Narrative orientation". However, the other two dimensions are specialized to the university discourse domain: Dimension 2 is interpreted as "Procedural vs. content-focused discourse". The co-occurring "procedural" features include modals, causative verbs, second person pronouns, and verbs of desire + *to*-clause; these features are especially common in classroom management talk, course syllabi, and other institutional writing. The complementary "content-focused" features include rare nouns, rare adjectives, and simple occurrence verbs; these co-occurring features are typical of textbooks, and especially common in natural science textbooks. Dimension 4, interpreted as "Academic stance", consists of features like stance adverbials (factual, attitudinal, likelihood) and stance nouns + *that*-clause; classroom teaching and classroom management talk is especially marked on this dimension.

A final example comes from Biber's (2008) MD analysis of conversational text types, which identified a dimension of "stance-focused vs. context-focused discourse". Stance focused conversational texts were marked by the co-occurrence of *that*-deletions, mental verbs, factual verbs + *that*-clause, likelihood verbs + *that*-clause, likelihood adverbs, etc. In contrast, context-focused texts had high frequencies of nouns and WH-questions, used to inquire about past events or future plans. The text type analysis identified different sets of conversations characterized by one or the other of these two extremes.

In sum, corpus-driven MD studies of English registers have uncovered both surprising similarities and notable differences in the underlying dimensions of variation. Two parameters seem to be fundamentally important, regardless of the discourse domain: a dimension associated with informational focus vs. (inter)personal focus, and a dimension associated with narrative discourse. At the same time, these MD studies have uncovered dimensions particular to the communicative functions and priorities of each different domain of use.

These same general patterns have emerged from MD studies of languages other than English, including Nukulaelae Tuvaluan (Besnier 1988); Korean (Kim and Biber 1994); Somali (Biber and Hared 1992; 1994); Taiwanese (Jang 1998); Spanish (Biber et al. 2006; Biber and Tracy-Ventura 2007; Parodi 2007); Czech (Kodytek 2008), and Dagbani (Purvis 2008). Taken together, these studies provide the first comprehensive investigations of register variation in non-western languages.

Biber (1995) synthesizes several of these studies to investigate the extent to which the underlying dimensions of variation and the relations among registers are configured in similar ways across languages. These languages show striking similarities in their basic patterns of register variation, as reflected by:

- the co-occurring linguistic features that define the dimensions of variation in each language;
- the functional considerations represented by those dimensions;
- the linguistic/functional relations among analogous registers.

For example, similar to the full MD analyses of English, these MD studies have all identified dimensions associated with informational vs. (inter)personal purposes, and with narrative discourse.

At the same time, each of these MD analyses have identified dimensions that are unique to a language, reflecting the particular communicative priorities of that language and culture. For example, the MD analysis of Somali identified a dimension interpreted as "Distanced, directive interaction", represented by optative clauses, first and second person pronouns, directional pre-verbal particles, and other case particles. Only one register is especially marked for the frequent use of these co-occurring features in Somali: personal letters. This dimension reflects the particular communicative priorities of personal letters in Somali, which are typically interactive as well as explicitly directive.

The cross-linguistic comparisons further show that languages as diverse as English and Somali have undergone similar patterns of historical evolution following the introduction of written registers. For example, specialist written registers in both languages have evolved over time to styles with an increasingly dense use of noun phrase modification. Historical shifts in the use of dependent clauses is also surprising: in both languages, certain types of clausal embedding—especially complement clauses—turn out to be associated with spoken registers rather than written registers.

These synchronic and diachronic similarities raise the possibility of universals of register variation. Synchronically, such universals reflect the operation of underlying form/function associations tied to basic aspects of human communication; and diachronically, such universals relate to the historical development of written registers in response to the pressures of modernization and language adaptation.

8.4 Conclusion

The present chapter has illustrated how corpus analysis contributes to the description of language use, in many cases allowing us to think about language patterns in fundamentally new ways. Corpus-based analyses are the most traditional, employing the grammatical categories recognized by other linguistic theories but investigating their patterns of variation and use empirically. Such analyses have shown repeatedly that our intuitions about the patterns of use are often inaccurate, although the patterns themselves are highly systematic and explainable in functional terms.

Corpus-driven approaches are even more innovative, using corpus analysis to uncover linguistic constructs that are not recognized by traditional linguistic theories. Here again, corpus analyses have uncovered strong, systematic patterns of use, but even in this case the underlying constructs had not been anticipated by earlier theoretical frameworks.

In sum, corpus investigations show that our intuitions as linguists are not adequate for the task of identifying and characterizing linguistic phenomena relating to language use. Rather, corpus analysis has shown that language use is patterned much more extensively, and in much more complex ways, than previously anticipated.

This page intentionally left blank

CHAPTER 9

DEFAULT SEMANTICS

KASIA M. JASZCZOLT

9.1 Default semantics and contextualism

Since the rise of radical pragmatics in the late 1970s, semantics has begun to grow to include not only the study of the meaning of the *sentence* but also those aspects of meaning intended by the author (speaker, writer) of this sentence which transform sentence meaning into the *speaker's intended, explicit meaning*, or *what is said*. To mention a few landmarks, Grice (1978) observed that pragmatic processes of disambiguation and reference assignment to indexical expressions sometimes contribute to what is said and therefore to the semantic representation of the sentence, allowing for an analysis of the proposition in truth-conditional terms. Kempson (1975; 1979) and Atlas (1977; 1979) proposed that semantic ambiguity of negation in English is better conceptualized as semantic underdetermination, where the logical form underspecified as to the scope of the negation operator is further enriched through pragmatic inference. This blurring of the boundary between semantics and pragmatics opened up an opportunity to reconsider the object of study of truth-conditional analysis. As a result, truth conditions became applied to a representation of *utterance meaning* which corresponds to the logical form of the sentence that can be enriched, or, to use a more general term, further *developed* (Sperber and Wilson 1986; 1995; Carston 1988; 2002) or *modulated* (Recanati 2004*a*; 2005) by some pragmatic additions such as the result of pragmatic inference from the context, and, on some accounts, non-inferential, automatically added pragmatic defaults from salient presumed scenarios. As a result, some of the meanings which Grice classified as implicit became reallocated to the explicit, truth-conditional content of utterances. This view belongs to the general orientation called *contextualism* (Preyer and Peter 2005). On its strong version, the propositional, truth-evaluable level of meaning is always subjected to such pragmatic embellishments and these embellishments are not dictated by the syntax or lexicon. According to weaker versions, such interaction of semantic and pragmatic output is not omnipresent but merely possible. While the contextualist orientation is not the only currently pursued construal of the semantics/pragmatics boundary, with minimalist semantics and various hybrid models making their way in (see Jaszczolt forthcoming a, b for an overview), it is arguably a successful way of representing Gricean, intended meanings, with an opportunity to provide a formal model of utterance meaning thanks to appropriating the tool of truth conditions on the "pragmaticky" side of the divide.

Default Semantics (Jaszczolt, e.g., 2005; 2009; henceforth DS) sits comfortably in the contextualist camp and in its radical flank, but also goes significantly beyond some of its assumptions. Its objective is to model utterance meaning as intended by the Model Speaker and recovered by the Model Addressee.[1] The constructs of model interlocutors are adopted in recognition of the view shared by Grice and the neo-Griceans that a theory of meaning should focus on general mechanisms that underlie the composition of meaning in conversation, including those mechanisms that make use of conventions, heuristics pertaining to rational human behavior, which explain shortcuts that addressees take through laborious inferential process of meaning recovery. These shortcuts are facilitated by standard, assumed scenarios and assumptions about human mental processes. In short, where context and inference need not be employed, they do not figure in the construction of meaning. Where DS goes beyond contextualism is its understanding of the interaction between the logical form of the uttered sentence and the kind of meaning that the theory of meaning has to represent. Unlike other contextualist accounts, it does not recognize the level of meaning at which the logical form is pragmatically developed/modulated as a real, interesting, and cognitively justified construct. To do so would be to assume that syntax plays a privileged role among various carriers of information and that the syntax/pragmatics interaction is confined to

[1] The question "whose meaning" a theory of discourse meaning should model is subject to ongoing discussions which frequently result in conflicting interpretations of Grice's Cooperative Principle, allowing it a normative or intensional, speaker-oriented interpretation (see Saul 2002 and Davis 1998, 2007 respectively). Post-Gricean approaches select the speaker's or the addressee's perspective (see Levinson 2000 and Sperber and Wilson 1995 respectively, with the proviso that Levinson's theory can also be read as normative). DS adopts the normative perspective for modeling utterance meaning, in the sense of behavioral norms pertaining to rational communicative behavior. See Jaszczolt 2005.

pragmatic additions, embellishments, or "developments" of the output of syntactic processing. If semantic theory is to model intended meaning understood as the most salient information conveyed by the Model Speaker, there does not seem to be any reason to impose this constraint. While pragmatic processing may develop the logical form in many cases, there is no reason to impose this requirement on *all* cases of communication. In fact, there is overwhelming empirical evidence that to do so is a mistake on the part of contextualists.

9.2 Primary meaning without the syntactic constraint

Grice proposed a fairly clear-cut distinction between what is said as the truth-conditional aspect of utterance meaning (his meaning$_{\text{nn}}$) on the one hand, and a battery of implicatures on the other. To repeat, what is said was identified with what was explicitly uttered, allowing also for some minimal help from pragmatics in establishing the propositional content in the form of reference assignment to indexical expressions and disambiguation. However, it can be easily observed that speakers don't always communicate their main, most salient messages through what is explicitly uttered. They also communicate them through what is implicated. In recent years Grice's "what is said" has been heavily criticized as too restrictive in its requirement that *saying* something must entail that the speaker *means* it and as a result as leading to problems with nonliteral use such as metaphor. If the speaker doesn't "say" but, as Grice suggests, merely "makes as if to say" that Tom is a slimy snake, then how can an implicature be produced? Implicatures are the result of saying something, not of "making as if to say". It is clear that the concept of saying should be relaxed in order to dispose of the awkward solution of making as if to say.

Although metaphor has now been revindicated as explicit rather than implicit content (Carston 2002), this takes us only part-way. If what is said can contain shifts from standard concepts such as SNAKE to newly constructed concepts fit for the particular context, then why can't it contain other meanings which are equally basic as intended content of the speaker's utterance? Why, for example, when a mother replies as in the exchange (1) below, what is said is to be represented as A rather than as B?

(1) Child: Can I go punting?
Mother: You are too small.
(A) The child is too small to go punting.
(B) The child can't go punting.

Similarly, in the celebrated example adapted from Bach (1994), the mother's response seems to primarily convey (B_1), (B_2), or a similar comforting statement. Contextualists, however, represent it as (A).

(2) Situation: A little boy cuts his finger and cries.
Mother: You are not going to die.

(A) The boy is not going to die from the cut.
(B_1) There is nothing to worry about.
(B_2) It's not a big deal.

DS takes as its object of semantic representation the primary, salient, intended meanings and hence allows for the B interpretations to be modeled. It is well documented that interlocutors frequently communicate their main intended content through a proposition which is not syntactically restricted in the way A interpretations are. In other words, the representation of the primary meaning need not be isomorphic with the representation of the uttered sentence. Further, it need not be isomorphic with any development of that syntactic form and hence need not constitute an enrichment or modulation of the proposition expressed in the sentence—if, indeed, the sentence happens to express a full proposition. Instead, it is quite common to convey the main message through an implicature.

Now, according to the definition widely adopted in contextualism, implicatures have semantic representations which are syntactically independent of the logical form of the uttered sentence. While the criteria by which the explicit and the implicit content of the speaker's utterance are a contentious matter (see Carston 1988, 1998; Recanati 1989, 2004*a*), there is widespread agreement that embellishments of the logical form of the sentence guarantee that we are talking about the explicit content. On the other hand, a proposition with an independent logical form—independent in virtue of entailment, or a psychological criterion of functioning as a separate premise in reasoning, or finally in virtue of being a wastebasket of communicated thoughts which do not qualify as *the* explicit content—is always an implicature. In this way, Grice's pool of implicata is reduced there to only those messages which do not develop the syntactic structure of the uttered sentence. In DS, we see no reason for this classification. While it is indeed sometimes the case that the main communicated message corresponds to such an enriched proposition, it is also frequently the case that the main communicated message corresponds to the bare sentence or to an altogether different sentence as in the B cases above. In DS, the syntactic constraint of post-Gricean contextualism is rejected. Instead, we argue that the kind of meaning that is modeled in the theory of meaning is the *primary meaning*. The primary meaning is the main message intended by the Model Speaker and recovered by the Model Addressee and it becomes the primary object of semantic analysis independently of its relation to the syntactic form of the uttered sentence. This meaning is construed on contextualist principles. We have a truth-conditional analysis of the representation that mixes semantic and pragmatic sources of information. In this construal we are merely taking one step beyond the post-Gricean contextualist approaches: we retain the Gricean ideology of intension-based meaning$_{nn}$ and reject the syntactic constraint on its content. There is ample experimental evidence that the primary meaning (or main meaning, explicit meaning, what is said, or whatever intuitive labels we want to assign to it) frequently corresponds to implicatures (Nicolle and Clark 1999; Pitts 2005; Sysoeva and Jaszczolt 2007 and forthcoming). The assumption followed in DS is that the object of study of a theory of meaning should reflect this fact of conversation and give a representation, and also a formal account, of precisely this intended, intuitive, most salient, and most important message.

9.3 Merger representation and its contributing sources

In DS, primary meanings are modeled as the so-called *merger representations*. The name reflects the important tenet of the theory that all sources of information about meaning provide constituent ingredients for the final representation. The outputs of these sources merge and all the outputs are treated on an equal footing, without giving priority to any of the sources. By the same token, the logical form of the sentence is not given priority over any other information. The syntactic constraint discussed in section 9.2 is abandoned. It is assumed that merger representations have the status of mental representations. They have a compositional structure: they are proposition-like, truth-conditionally evaluable constructs, integrating information coming from various sources that interacts according to the principles established by the intentional character of discourse. In the revised version of DS, there are five sources of such information:

(i) world knowledge (WK);
(ii) world meaning and sentence structure (WS);
(iii) situation of discourse (SD);
(iv) properties of the human inferential system (IS);
(v) stereotypes and presumptions about society and culture (SC).[2]

The source WK pertains to information about the laws governing the physical world, such as that leading to the interpretation of *and* as *and as a result* in (3b).

(3) a. The temperature fell below −10 degrees Celsius and the lake froze.
b. The temperature fell below −10 degrees Celsius and as a result the lake froze.

[2] The revised version of DS was first discussed in Jaszczolt 2009. This summary makes use of some ideas introduced there.

WS stands for information from the lexicon and syntax. SD pertains to the context provided by the situation in which the utterance is issued, including the participants, location, time, and co-text. The source IS stands for the structure and operations of the human brain that are responsible for the emergence of standard, default interpretations of certain types of expressions, unless the addressee has evidence that this standard interpretation is not intended. IS is responsible, for example, for the default referential as opposed to attributive reading of definite descriptions as in (4b).

(4) a. The author of *Cloud Atlas* has breathtaking sensitivity and imagination.
b. David Mitchell has breathtaking sensitivity and imagination.

It is also responsible for the default *de re*, as opposed to *de dicto*, interpretation of intensional constructions such as belief reports (see Jaszczolt 2005, 2007*a*). Finally, SC is a source of information which is not specific to the situation of discourse but instead is shared across different scenarios in a socio-cultural and linguistic community. It is responsible, for example, for the interpretation of (5a) as (5b).

(5) a. A Botticelli was stolen from the Uffizi last week.
b. A painting by Botticelli was stolen from the Uffizi Gallery in Florence last week.

The sources are presented in Figure 9.1. Merger representation, the result of the interaction of information they provide, is referred to as Σ in that this symbol reflects the *summation* of content coming from WK, WS, SD, IS, and SC that takes place at this level of representation.

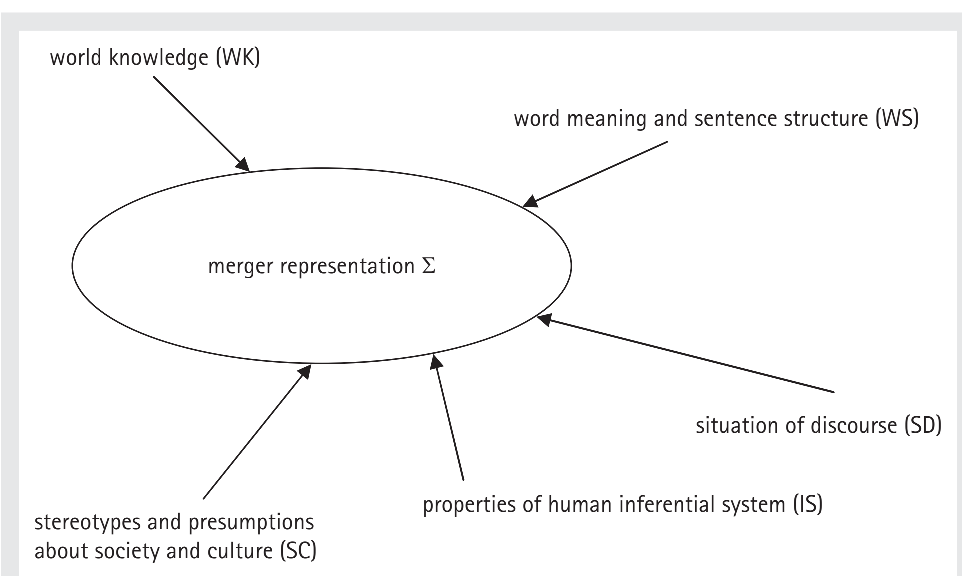


Figure 9.1. Sources of information contributing to a merger representation Σ

To sum up, Σ is the result of the interaction of information about the primary meaning as intended by the Model Speaker and recovered by the Model Addressee—an interaction that draws on the five sources of meaning distinguished here as (i)–(v) and presented in Figure 9.1. All sources of meaning operate on an equal footing. The primary meaning obtains the representation that is provided solely by a semantic composition that proceeds according to the heuristics of rational conversational behavior, modeled on methodologically adequate and testable principles of economy of agent input and maximization of obtained information (see, for example, Horn 1988; Levinson 1987*a*, 2000; Hawkins 2004*a*).

In *Default Semantics* (Jaszczolt 2005), I identified four sources of information about meaning that make up merger representations: word meaning and sentence structure; conscious pragmatic inference; social and cultural defaults; and cognitive defaults. The difference between these sources and the revised list of sources (i)–(v), presented also in Figure 9.1, pertains to the criteria by which they are individuated. In Figure 9.1, I used qualitative labels such as culture, society, physical laws, or context. In the 2005 version, I used processing criteria. I tried to demonstrate how the type of processing is linked to the particular source. As a result, the types of information were matched with the types of processing in order to produce the four categories of the 2005 version of DS. For example, the fact that shared knowledge of Italian Renaissance painting, such as that of Botticelli in example (5a), may activate automatic and subconscious interpretation in (5b) explains the separation of the category labeled as "social and cultural defaults" from the category "conscious pragmatic inference". On the other hand, the referential identification of "Larry" in (6) as Larry Horn is more likely to be the result of a conscious inferential process.

(6) Larry's account of the history of negation is truly magnificent.

In the revised version (Jaszczolt 2009), sources are differentiated from processing mechanisms. The model of sources of information can be mapped onto *types of processes* that produce both the merger representation Σ of the primary meaning and the additional (secondary) meanings. The labels "primary meaning" (pm) and "secondary meaning" (sm) are preferred to the traditional "what is said" and "what is implicated" in that, to repeat, the primary meaning can be a Gricean implicature—the most salient intended meaning need not correspond to the uttered sentence. These labels are also preferred to the terms "Stage I" and "Stage II" used in the 2005 version in that the term "stages" is conducive to interpreting the model in terms of the temporal sequence of processing, whereas there is empirical evidence that the processing of primary and secondary meanings does not have to proceed in this particular order. Processing of some secondary meanings may be necessary en route to recovering the primary message. The revised *processing model* of utterance interpretation in DS is given in Figure 9.2.

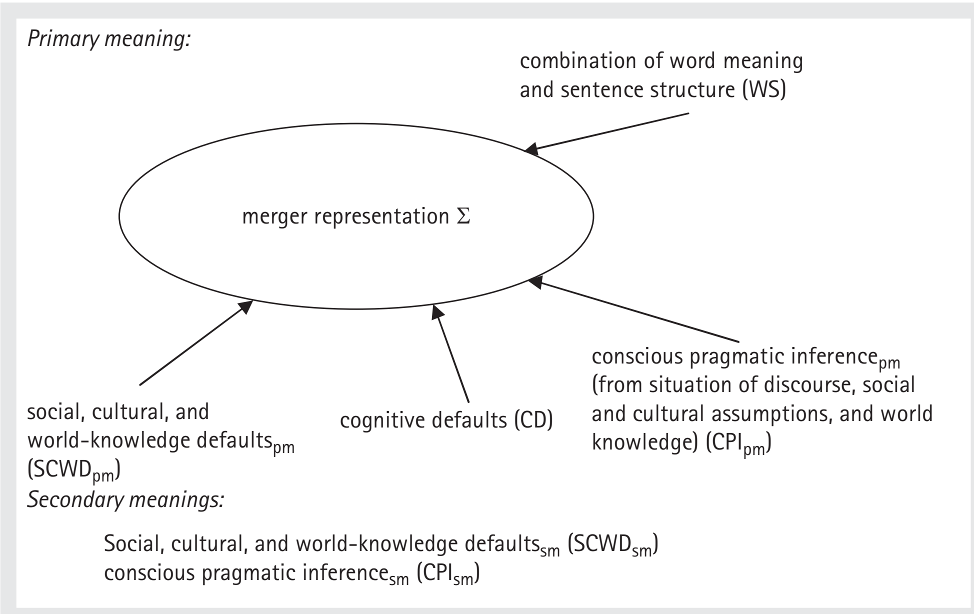


Figure 9.2. Utterance interpretation according to the *processing model* of the revised version of Default Semantics

The mapping between the "sources" model and the "processing" model is as follows. World knowledge (WK) and stereotypes and presumptions about society and culture (SC) can lead to automatic, default interpretations of SCWD kind, as in example (5) above, but also to consciously, inferentially reached ones (CPI), as in (6). Word meaning and sentence structure (WS) constitute both a source and a type of processing, according to the modularity assumption discussed in the following section. They simply produce the logical form of the uttered sentence. Situation of discourse (SD) triggers CPI. Properties of human inferential system (IS) result in a kind of default interpretations that pertain to the structure and operations of the brain, and hence there is a one-to-one correspondence between source IS and a process that produces CD. For building merger representations DS makes use of the processing model and it indexes the components of Σ with a subscript standing for the type of processing, as is exemplified in the selected applications in section 9.7.

Next, as is evident from the model in Figure 9.2, social and cultural stereotypes can play a part in the construction of the primary meaning as well as the secondary meanings. Similarly, context-driven conscious inferences can contribute both to the primary meaning and the secondary meanings. They are active in the recovery of additional meanings intended by the Model Speaker and recovered by the Model Addressee.

The question of what counts as effortful processing (CPI) vis-à-vis automatic utilization of knowledge of culture and society (SCWD) is a difficult and currently unresolved one. While it is widely accepted that utterance interpretation makes use of automatic, default interpretations, which are assumed and which figure as salient and strong interpretative probabilities unless the context dictates otherwise, it is not possible to tell where the boundary lies. In fact, it may never be possible to make reliable generalizations on this matter, due to the interpersonal differences in assumed common ground. While "Botticelli" may trigger for most people the interpretation in (5b) by means of SCWD, this default status cannot be universally assumed for all interlocutors. Substituting names of lesser known artists may give even less default-oriented results. Example (7) gives the flavor of the extent of the problem. While in the scenario of an interview with a film star it can be assumed that "Leonardo" refers to a young but widely known actor, Leonardo DiCaprio, assuming this as a cultural default for the interviewer would be too strong. In fact, assuming it as a salient referent reached as a result of conscious pragmatic inference is proven to be equally wrong. The scriptwriter obtains the humorous effect through a mismatch of the very salient referent intended by the interviewed young actress (B) and the one which is recovered by the interviewer (A) unexpectedly for the viewers of the film. The assumption that the reference assignment as Leonardo DiCaprio is a result of SCWD rather than CPI makes the situation even more striking.

(7) A: So, is this your first film?
B: No, it's my twenty-second.
A: Any favorites among the twenty-two?
B: Working with Leonardo.
A: da Vinci?
B: DiCaprio.
A: Of course. And is he your favorite Italian director?
(Richard Curtis, *Notting Hill*, 1999)

The problem is this. While in the analysis of a particular example we may not be able to univocally allocate the interpretation to SCWD or CPI, there is a clear need to distinguish the two kinds of processes: the conscious, inferential one and the automatic, subdoxastic one, i.e., operating below the level of thoughts of which we are conscious.[3] DS finds a solution to this uncertainty in that merger representations are representations produced by the *Model* Addressee and contain the meanings intended by the *Model* Speaker. On this level of theory construction, it is quite plausible to make assumptions about the sources. Even if they do not correspond to those utilized by actual interlocutors on every occasion, the adequacy of the

[3] I have surveyed various theories of meaning that make use of default interpretations in several other works. See for example Jaszczolt 2006 for a comprehensive encyclopedic account.

framework is not compromised. It also has to be pointed out that any model of utterance interpretation which, like DS, or Levinson's (2000) presumptive meanings, retains the common intuition that the primary meaning is built *both* out of automatic, associative, unreflective components and conscious, inferential ones, has a significant advantage over restrictive and highly implausible accounts on which all such components of primary meaning are inferential or all are associative. Allowing for some differences among the accounts compared here that pertain to the acceptance or rejection of the syntactic constraint, the argument goes as follows. If, like Carston (e.g., 2002, 2007), one argues for only inferential enrichments, some of those "inferences" will have to be dubbed subconscious and spontaneous. If, on the other hand, one opts for only associative enrichments like Recanati (e.g., 2002, 2004, 2007), one has to qualify this view by saying that some of them may be open to retrospection and be conscious in this sense. A common-sense "inference plus defaults" account of DS is clearly superior.

9.4 Compositionality of primary meanings

The next issue to be addressed is the compositionality of Σ. It is well known that the requirement of compositionality of meaning is the stumbling block of all extant semantic theories. It has been one of the greatest challenges of semantic theory to provide a compositional account of such intensional contexts as propositional attitude reports, modal expressions, and constructions with temporal adverbials. However, as was extensively argued in the recent contextualist literature (Recanati 2004*a*; Jaszczolt 2005), compositionality need not be imposed on the level of the output of syntax alone. While it is arguably the case that a theory of meaning which is not compositional in some sense is inconceivable and the methodological requirement of compositionality has to be present on one level or the other, compositionality need not be as strict as the traditional truth-conditional semantics requires. For Schiffer (e.g., 1991, 1994, 2003), compositional semantics is not a *sine qua non* condition; composition of meaning may simply reflect compositional reality. In other words, meaning supervenes on the structure of the world. For Recanati, compositionality belongs to the level of truth-conditional pragmatics—the level of enriched, modulated propositions. DS follows this approach and places the methodological requirement of compositionality on the representation of *utterance* meaning rather than *sentence* meaning. Recanati (2004*a*: 138) calls it the Pragmatic Composition view and an "interactionist", "Gestaltist" approach to compositionality (ibid.: 132). But a terminological clarification is necessary here. While for Recanati this view is properly called *truth-conditional pragmatics*, DS retains the name *truth-conditional semantics*, allowing for semantics to be construed in the contextualist way—or, even, a radical contextualist way since the syntactic constraint is rejected there. Composition of meaning is not dictated by the syntactic form of the uttered sentence but rather by the intended meaning of the speaker. Compositionality is predicated of merger representations—the Σs of primary meanings of utterances, as intended by Model Speakers and recovered by Model Addressees. When the task of semantics is to produce a merger representation, the semantic composition is necessarily largely "pragmaticky": semantic composition means composing merger representations Σ.

Now, like its parent theory, Discourse Representation Theory (henceforth DRT, e.g., Kamp and Reyle 1993) on which DS is loosely modeled, DS regards formalization as subordinate to the overall goal of constructing representations of discourse. These representations are assumed to be cognitively real and reflect mental processing of natural language utterances (Hamm et al. 2006). It is also in agreement with Jackendoff (2002; Culicover and Jackendoff 2005), propounding that *conceptual semantics* is a superordinate objective: it makes use of formal methods but is not constrained by their limitations. Pragmatic compositionality also shares the overall orientation with some new developments in the philosophy of language where the problem of substitutivity in intensional contexts is progressively further removed from the problem of substitution of coreferential expressions in sentences. For example, Pelczar (2004, 2007) proposes that the fact that a speaker may hold a certain belief about water, say, that it is in short supply, but fail to hold a belief that corresponds to "H_2O is in short supply" should not be approached as a problem with the properties of the objects of belief but as a problem with the *attitude of believing* itself. This means that, rather than complicating the theory of the objects of thought à la Kaplan, Perry, or Schiffer, he proposes that there is one single object of thought, referred to as "water", "H_2O", and some other contextually salient labels, but belief relation itself is represented as a context-dependent, indexical predicate whose content depends on the features of the context such as the topic of conversation, conceptual background of the interlocutors, or the discursive history (co-text). The problem with compositionality in belief reports is therefore solved by appeal to the representation of context which makes the belief itself, and the belief expression, indexical. Pelczar calls this view *formal pragmatics* in that literal content is allowed to depend on contextual factors, including norms and maxims of conversation. He adds that this construal need not violate the principle of compositionality: in a compositional theory of meaning, expressions which enter into this composition have themselves contextually determined contents (see Pelczar 2004: 71).[4] DS goes a little further. In liberating merger representations from the syntactic constraint, it

[4] See also Predelli 2005*a*, *b* on the standard truth-conditional semantics being sufficiently "contextualist".

brings truth-conditional methods closer to cognitive, conceptual analyses, arguably to the mutual benefit of conceptual and truth-conditional aspects of the analysis of meaning.

To sum up, while compositionality is a necessary prerequisite for any theory of meaning and need not be questioned, principally because one has nothing better to hold on to instead, compositionality should not be seen as a methodological requirement on the syntax and semantics of *sentences*. Such a narrow view of compositionality has been proven to complicate formal methods in order to fit natural language into the mold of formal languages of deductive logic. Instead, DS agrees with Jackendoff (2002: 293) that there is no "strictly linguistic meaning" and that constructing mental representations of discourse is the fundamental objective of a theory of meaning. This more pragmatic approach to compositionality permeates some recent accounts which are engendered by the disillusionment with the strict methodological requirements of post-Montagovian theories. Dowty (2007), for example, suggests that compositionality is not a "yes-no question" but rather a "how-question" and belongs in the empirical domain of facts:

> I propose that we let the term NATURAL LANGUAGE COMPOSITIONALITY refer to *whatever strategies and principles we discover that natural languages actually do employ to derive the meanings of sentences, on the basis of whatever aspects of syntax and whatever additional information (if any) research shows that they do in fact depend on*. Since we do not know what all those are, we do not at this point know what "natural language compositionality" is really like; it is our goal to figure that out by linguistic investigation. Under this revised terminology, there can be no such things as "counterexamples to compositionality", but there will surely be counterexamples to many particular hypotheses we contemplate as to the form that it takes. (Dowty 2007: 27)

DS endorses this view. Compositionality retains the status of the methodological requirement in that it is assumed that it has to be *discovered* rather than to be *shown whether* it is the case. The main question becomes, what principles are responsible for the compositional, and hence calculable in the sense of "predictable", character of the *entire system* of human communication? In other words, while non-compositional theory of meaning is not considered (*pace* Schiffer's 1991, 1994, 2003 compositional supervenience but non-compositional semantics), "kicking compositionality up", so to speak, from the level of pure syntax to a multidimensional, merged, interactive, representation of discourse meaning is offered as a plausible option.[5] This level is the level of a merger representation Σ and therefore compositionality is predicated of the result of the interaction of information provided in the processes summarized in Figure 9.2 above. It is now a task for future experimental

[5] Not only need compositionality not be constrained to sentences, it need not even be constrained to approaches espousing truth conditions and reference. For a deflationist version of compositionality see Horwich 2005.

projects to pave the way toward capturing this interaction of sources more formally as an algorithm for merging the outputs.

9.5 Incremental processing

Another pertinent and as yet unresolved problem in Gricean pragmatics concerns the global vs. local character of the additions to the logical form of the uttered sentence. On Grice's (1978) original proposal, what is said consisted of the truth-conditional component of meaning$_{nn}$. Reference assignment and disambiguation operated within the boundaries set by syntax. Indexicals provided syntactic slots for referents. Lexical and syntactic ambiguity was a self-evident output of the grammatical analysis and hence disambiguation was also internally motivated by the system. Next, implicatures of various kinds were constructed on the basis of the inference from the proposition that corresponded to the sentence, disambiguated, and referentially pinned down as required. All implicatures were dependent on the meaning of the uttered sentence which was processed first. In other words, they were *post-propositional*. Even the enrichment of sentential connectives such as that from *and* to *and therefore* or *and then* in (8b), for which the sub-maxim of manner, "be orderly", was said to be responsible, was construed as such a global, post-propositional process of implicature recovery.

(8) a. Mary finished marking the scripts and went to the cinema.
b. Mary finished marking the scripts and then went to the cinema.

However, with the growing emphasis on the psychology of utterance processing in post-Gricean pragmatics, the global character of such additions began to be questioned. The most radical form of this reaction is arguably Levinson's (2000) theory of Generalized Conversational Implicature (henceforth GCI) and his so-called presumptive meanings. While Levinson remains close to the spirit of Grice in emphasizing the context-free character of some enrichments, and thereby retaining the contentious category of generalized implicature, he also emphasizes the incremental character of discourse processing and proposes local enrichments, triggered by parts of the sentence such as phrases, words, and sometimes morphemes. (9)–(11) below exemplify Levinson's local GCIs, where the symbol "+ >" stands for "conversationally implicates".

(9)	Some of the boys came.	+ > "not all".
(10)	Possibly, there's life on Mars.	+ > "not certainly".
(11)	If John comes, I'll go.	+ > "maybe he will, maybe he won't" (Levinson 2000: 36–7).

These interpretations are explained by the Q-heuristic, "What isn't said, isn't": where a stronger expression is available but was not used, it can be inferred that using it would lead to falsehood. Examples (9)–(11) allegedly demonstrate that the presumptive meaning arises as soon as the smallest relevant item is processed. In other words, it is not the case that the proposition "Some of the boys came" leads to the interpretation "Not all of the boys came", as was the case on Grice's original construal, but rather it is the word "some" that triggers "not all". However, the problem is that the approach seems to be self-defeating: the more we try to obey the principles of rational communicative behavior, the more we contradict them. Levinson argues that it is the "bottleneck of communication" that makes us say less and infer more. Articulation is costly and slow, processing is cheap and fast. Addressees tend to go beyond the words to the standard, salient, default interpretation at the first encountered opportunity. But the sooner this "defaulting" takes place, the greater the risk that is it incorrect and has to be revoked. In other words, if, say, "some" is resolved as "some but not all" as soon as the word is processed, then all cases that follow the patterns of (9a) or (9b) below will necessitate the cancellation of the default interpretation.

(9) a. Some of the boys came. In fact, all of them did.
b. Some, and in fact all, of the boys came.

It goes without saying that default interpretations are defeasible. But their cancellation has to be restricted to cases where it can be plausibly predicted to be happening. In short, making defaults too local makes them too costly to be true. While the existence of shortcuts in reasoning is not disputed, the exact character and constitution of these shortcuts is still largely unknown. In this section I look at their constitution, and mainly at their length. The characteristics of default interpretations vis-à-vis inferential ones are taken up in section 9.6.

Locality is even more pronounced in examples (12)–(15).

(12) bread knife +> knife used for cutting bread
kitchen knife +> knife used for preparing food, e.g., chopping
steel knife +> knife made of steel

(13) a secretary +> female one

(14) a road +> hard-surfaced one

(15) I don't like garlic. +> I dislike garlic.
[triggered locally by "don't like", KJ] (adapted from Levinson 2000: 37–8).

These cases are explained by his I-heuristic, "What is expressed simply is stereotypically exemplified". Locality so construed is not without problems. (12) contains three standard compounds, in spite of the orthographic conventions which make them written as separate words, and it is not at all clear that compounds can be subsumed under the category of pragmatic enrichment. And, naturally, if they are not cases of GCI, they do not constitute a suitable argument for locality. Examples (13) and (14) are also problematic. It is a commonly shared intuition that "secretary" in "the Prime Minister's secretary" does not default to "female". "Road" is a lexical item that comes with the conceptual content that can be variously construed as a prototype, definition, set of features, depending on the adopted theory of word meaning, and that contains a concept of hard, leveled, drivable, or at least walkable, surface of some sort. Just as in the case of compounds in (12), it seems that instead of defaulting to "local" enrichment we have here a simple lexical content of the word "road" *tout court*. Next, in (15), the *neg*-raising is rather hastily classified as local. It seems at least plausible that the "dislike" interpretation arises because of the content of the sentence, the state of not being fond of garlic, rather than as a result of the strengthening of "not like" alone.

In short, the more "local" the enrichments, the higher the likelihood that they have to be canceled later on in discourse when more information becomes available. Frequent cancellation is not a satisfactory feature of defaults in that it is costly. So, it seems that one has to opt for a solution that combines the fact of the incremental nature of processing with the fact that cancellation is not welcome. Adopting Grice's original post-propositional, "global" enrichments would mean that the cost and frequency of cancellations is substantially reduced. But then, developing this line of argument by allowing default interpretations to operate on units even larger than a single proposition, when this is appropriate to construe them in this way, would take us even closer to an adequate model of utterance interpretation. We can conclude that default interpretations should be construed as operating on a unit that is adequate for the particular case at hand. By "adequate" we mean cognitive, psychological adequacy that can be corroborated by experimental investigations into the characteristics of discourse processing. Pragmatic inference should be construed as similarly flexible.

Default Semantics does not as yet have a satisfactory answer to the question as to how to model the locality of default interpretations. However, it has an interim solution that avoids the pitfalls of radical globalism as well as radical localism. To repeat, the desideratum is this:

Default and inferential interpretation have to be construed as operating on a unit that is adequate for the case at hand, ranging from a morpheme to the entire discourse.

We do not as yet have even a descriptive generalization pertaining to this desideratum. Therefore, until evidence is amalgamated and classified, DS proposes to stay close to the Gricean spirit and analyze all default and inferential meanings as if they were interacting with the proposition-like unit given in the uttered sentence. Naturally, sometimes the sentence will correspond to a full proposition, at other times it will not (see Bach 1994, 2004*a*, 2006*a*). This issue, albeit important, will not concern us at the moment. What matters is that DS makes a methodological move to model the meanings that come from the source ws in Figures 9.1 and 9.2 globally:

word meanings and structure are subjected to semantic composition first, as far as such is available, and then they interact with the remaining aspects of meaning. Psychological reality is as yet a goal to be aspired to. As the interim measure we stay close to the assumption that reflects a more economical, more rational way for the communicators to proceed.

Compared with Grice's globalism and Levinson's localism, DS displays another advantage. Cognitive, cultural, social, and world-knowledge defaults referred to in DS as CD and SCWD do not pertain to the enrichment of the logical form understood as the output of syntactic processing. Since the syntactic constraint is abandoned, they can also "override" it. In composing utterance meaning, the output of syntactic processing is not pragmatically enriched but instead all the sources of information about meaning are equal contributors to the merger representation Σ, the representation of the proposition pertaining to the primary meaning.[6] When the interaction among the outputs of different constituent processes is construed in this way, that is, when they all operate on an equal footing, the question of locality and globality of "additions" to the logical form has to be reformulated as the question of the interaction of WS with the remaining aspects of Σ. It is no longer "additions" that we are talking about but "properties of the interaction". And, since we are far from the complete algorithm for meaning construction, it comes as no surprise that we are far from knowing the properties of WS either. What we have is a model of the end product, so to speak, and a theory of what sources (Figure 9.1) and what types of information (Figure 9.2) contribute to the merger. This is a big step in itself from the by now traditional view of the "enrichment of the logical form" but there remains a lot to be done both on the front of the psychology of processing and the formalization of the merger.

9.6 Default vs. inferential components of Σ

In section 9.3, four types of information about utterance meaning were identified: combination of word meaning and sentence structure (WS), conscious pragmatic

[6] DS subscribes to what Bach (2006*a*) calls *propositionalism*, a view that the proper object of study of a theory of meaning is a proposition and that that proposition is recovered from the sentence and the context—in DS in the form of default meanings and pragmatic inference. Bach rejects this view and opts instead for what he calls *radical semantic minimalism* according to which semantics need not concern itself with propositions and need not deliver truth conditions (see also Bach 2004*a*). On different forms of minimalism in semantics see Jaszczolt 2007*b*.

inference (CPI), cognitive defaults (CD), and social, cultural, and world-knowledge defaults (SCWD). The WS source was discussed in section 9.5 in relation to the question of the local vs. global character of the interaction with the other sources that produces the merger representation Σ. The CPI source was introduced in detail in section 9.3 where it was emphasized that DS construes pragmatic inference as a conscious process. The pleonastic label "conscious inference" is retained there in order to make this fact clear vis-à-vis other post-Gricean approaches on which "inference" is a more widely construed concept (see Levinson 2000; Recanati 2004*a*, 2007; Carston 2007). Such inference is not necessarily deductive: DS also admits defeasible forms of processing, such as inductive and also abductive reasoning, the so-called default reasoning (Thomason 1997; Jaszczolt 2006). It draws on information available in the situation of discourse, social and cultural assumptions of the interlocutors, as well as their knowledge of the physical laws governing the world. Information which is not produced via such conscious inference is thereby non-inferential, automatic. It is also called subconscious, subdoxastic, and sometimes associative (Recanati 2004*a*, 2007). It was observed in section 9.3 that, although there is no disagreement in the field concerning the existence of such shortcuts through inference, the actual properties and membership of this category are far from being resolved. As I argued extensively elsewhere (Jaszczolt 2006), although default interpretations are almost universally recognized in accounts of discourse meaning, what various authors mean by "default" differs on at least the following fronts:

[1a] Defaults belong to competence.
vs.
[1b] Defaults belong to performance.

[2a] Defaults are context-independent.
vs.
[2b] Defaults can make use of contextual information.

[3a] Defaults are easily defeasible.
vs.
[3b] Defaults are not normally defeasible.

[4a] Defaults are a result of a subdoxastic, automatic process.
vs.
[4b] Defaults can sometimes involve conscious pragmatic inference.

[5a] Defaults are developments of the logical form of the uttered sentence.
vs.
[5b] Defaults need not enrich the logical form of the sentence but may override it.

[6a] Defaults can all be classified as one type of pragmatic process.
vs.
[6b] Defaults come from qualitatively different sources in utterance processing.

There is also disagreement concerning the following properties, to be discussed below:

[7a] Defaults are always based on a complete proposition.
vs.
[7b] Defaults can be "local", "sub-propositional", based on a word or a phrase.

[8a] Defaults necessarily arise quicker than non-default meanings. Hence they can be tested for experimentally by measuring the time of processing of the utterance.
vs.
[8b] Defaults do not necessarily arise quicker than non-default meanings because both types of meaning can be based on conscious, effortful inference. Hence, the existence of defaults cannot be tested experimentally by measuring the time of processing of the utterance.

Some of these properties are interrelated, some just tend to occur together, and some exclude each other. At the current stage of its development, DS tends to favor the following cluster:

[1a], in that merger representation Σ is construed as a semantic representation;

[2b], in that salient, short-circuited interpretations arise through repeated exposure to scenarios and to information about culture, society, and physical world;
[3b], since, as was argued extensively in section 9.5, frequent cancellation goes against the economy and thereby rationality of communicative behavior;
[4a], in virtue of the very nature of what constitutes a default interpretation, as was discussed above;
[5b], following the rejection of the syntactic constraint in DS, discussed in section 9.2;
[6b], in that default interpretations are classified in DS as (i) CD, pertaining to the source IS, and (ii) SCWD, where SCWD pertain to two sources: WK and SC (see Figures 9.1 and 9.2);
[7a], as a temporary methodological measure, recognizing the reality of [7b], as was argued in section 9.5;
and
[8a], logically following [4a] and hence in virtue of the very nature of what constitutes a default interpretation.

This completes the introduction to the principles and desiderata of DS. The following section presents some selected applications of the theory.

9.7 Selected applications

DS is still a theory in progress. Its origins date back to the early 1990s and to the questioning of the need for the stage of utterance interpretation which pertains to the development of the logical form of the sentence. This question was first expressed in Jaszczolt 1992 and further elaborated in 1999, and called the *Parsimony of Levels (POL) Principle*: levels of senses are not to be multiplied beyond necessity (Jaszczolt 1999: xix). The first applications of the theory were to definite descriptions, proper names, and belief reports and thereby to other non-factive propositional attitude constructions (e.g., Jaszczolt 1997, 1999), followed by a DS-theoretic account of negation and discourse connectives (Lee 2002). Subsequently, DS was used for a wide range of constructions and phenomena that are standardly considered as problematic for semantic theory: presupposition, sentential connectives, number terms, temporality, and modality (see, for example, Jaszczolt 2005), the latter also in a contrastive perspective (Srioutai 2004, 2006; Jaszczolt and Srioutai forthcoming). One of the current projects applies DS-theoretic analysis to a selected class of speech acts in Russian and English, also testing experimentally the validity of the rejection of the syntactic constraint (Sysoeva and Jaszczolt 2007 and forthcoming). Another recent project developed DS-theoretic representations of the future, the present, and the past, demonstrating their cross-linguistic application, and proposed a new concept of merger representation (Σ') that replaces the murky concept of event and functions as an object of the propositional operator of temporality. Temporality is analyzed there as derivable from the concept of epistemic modality (Jaszczolt 2009). In this section I exemplify the use of the DS-theoretic analysis in two semantic domains: that of (i) definite descriptions and (ii) temporality: the representation of the past, present and future.

9.7.1 Merger representations for definite descriptions

Referring in discourse is performed by means of two categories of expressions: the directly referring (type-referential) ones, and the ones whose referring function is facilitated by the context (the token-referential ones). The first category comprises ordinary proper names, some pronouns, including demonstratives, and demonstrative phrases. Token-referential expressions are normally instantiated by definite descriptions, that is, by definite noun phrases that are used to refer to an object (rather than, say, generically). However, definite descriptions do not fit neatly into this classification. As has been well known since Donnellan's seminal paper (1966), they can be used to refer to a particular, interpersonally identifiable individual, i.e., referentially, or they can be used to attribute a certain description to whoever fulfills it, i.e., attributively. Let us imagine a situation in which the speaker points at the cathedral Sagrada Família in Barcelona and utters (16).

(16) The architect of this church was an eccentric.

What the sentence means, after the strictly Gricean filling in of the demonstrative noun phrase with a referent, is that the architect of Sagrada Família was an eccentric. But the actual meaning recovered by the addressee will depend on the referential intention assigned to the speaker. This referential intention can be weaker or stronger and can render the attributive meaning in (16a) or the referential meaning in (16b) respectively.

(16) a. The architect of Sagrada Família (whoever he was) was an eccentric.
b. Antoni Gaudí was an eccentric.

In DS, the commonly asked question as to whether the referential/attributive distinction is a semantic or a pragmatic one need not be posed in that the radically contextualist orientation of the theory, and the interactive provenance of merger representations, render it meaningless. But the duality of use remains the fact of discourse and Σ has to reflect it. In fact, there is more than duality here when we approach the problem from the perspective of discourse processing. When the addressee is referentially mistaken and believes that it was, say, Simon Guggenheim who designed this cathedral, the "recovered" reading is as in (16c).

(16) c. Simon Guggenheim was an eccentric.

All of these possibilities of reading of (16) have to be captured in merger representations because they all pertain to possible outcomes of the processing. In DS, it is assumed that the referentially strongest interpretation in (16b) is obtained by means of the interaction of WS with CD: strong referentiality is founded on the strong intentionality of the relevant mental acts and therefore constitutes the default interpretation for referring expressions. This is summarized in the principles of DS called *Degrees of Intentions (DI)* and the *Primary Intention (PI)*, stating respectively that intentions and intentionality allow for various degrees, and that the primary role of intention in communication is to secure the referent of the speaker's utterance (Jaszczolt 1999: xix). In other words, the referential, and referentially accurate, interpretation in (16b) comes out as the default interpretation of the cognitive type (CD), secured by the properties of the human inferential system (IS) in that the intentionality of the mental act that corresponds to this proposition is the strongest, undiluted by the lack of information for identification as in (16a) or by a referential mistake as in (16c). The resulting Σ is given in Figure 9.3, where x stands for the discourse referent and the formulae below for discourse conditions—the dynamic equivalents of the logical forms, construed in the amended and extended language borrowed from DRT. The subscript after the square bracket stands for the type of information that takes part in the interaction producing Σ.

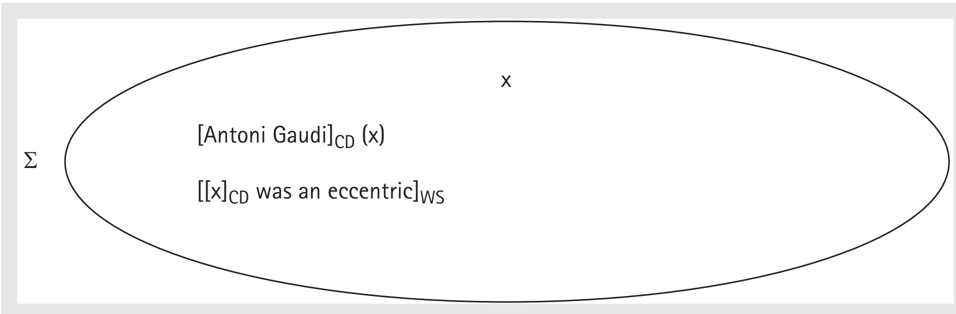


Figure 9.3. Merger representation for the default reading of example (16)

The reading with the referential mistake in (16c) comes next in that it exhibits a weaker form of intentionality and thereby a weaker referentiality. This reading owes its reference assignment to pragmatic inference and hence CPI_{pm} (see Figure 9.2), with the disclaimer carried forward from the earlier discussion that the assignment of the inferential or default route is at present somewhat speculative: while the distinction between inferential and automatic routes can hardly be contested, the allocation of particular cases to the categories has not as yet acquired a descriptive generalization. This reading is represented in Figure 9.4.

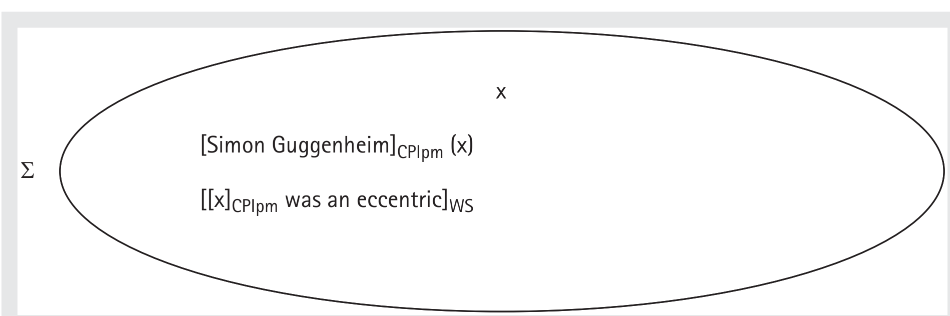


Figure 9.4. Merger representation for the referential mistake reading of example (16)

Finally, the attributive reading in (16a) pertains to the weakest referential intention and the weakest intentionality. It is represented in Figure 9.5. As in the case of the referential mistake, the merger representation is obtained through the interaction of the inferential (CPI_{pm}) identification of the referent and the WS. In the case of this reading, we also have to represent the composition of the phrase “the architect of this church”. “This church” is a demonstrative phrase in our example and hence a directly referring expression (type-referential). In the semantic representation we substitute the salient, correct referent, signaled as the proper name Sagrada Família, which is obtained as a cognitive default (CD). The definite description “the architect of Sagrada Família” is then composed with the help of WS (syntactic composition) and pragmatic inference (CPI_{pm}) in that this is the attributive, and hence non-referential, non-standard, inferentially obtained interpretation where referential intention with which the utterance was made and the corresponding intentionality of the mental act are both weaker as compared with the default case of (16b) and the mistaken referential (16c).

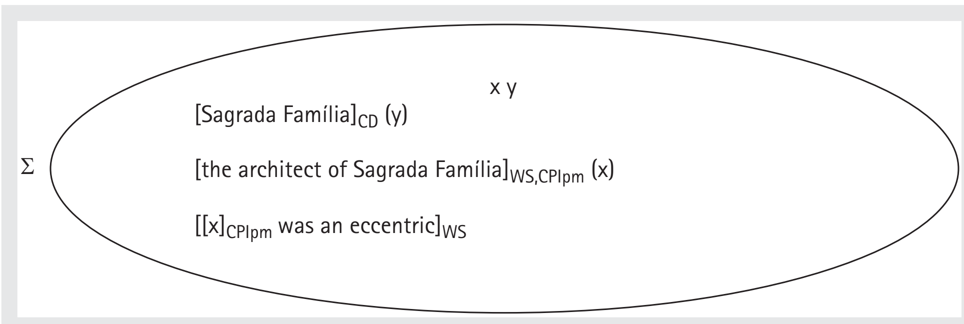


Figure 9.5. Merger representation for the attributive reading of example (16)

These merger representations have to be qualified as partial representations in that the temporality of the proposition is not represented. For the sake of clarity, we focused only on the representation of the referring expressions. The application of DS to the representation of the interlocutors' concept of time is the example discussed in the following section.

9.7.2 Merger representations of time

Referring to past, present, and future eventualities (events, states, processes) can be performed in a variety of ways. For example, in English, all of the expressions in (17a)–(17e) concern a future event.

(17) a. Lidia will play in a concert tomorrow evening.
 b. Lidia will be playing in a concert tomorrow evening.
 c. Lidia is going to play in a concert tomorrow evening.
 d. Lidia is playing in a concert tomorrow evening.
 e. Lidia plays in a concert tomorrow evening.

These future-time referring expressions vary somewhat in their semantic import. The expression *will* in (17a) and (17b), called in DS regular future, is attested to be undergoing remodalization: from the modal meaning of volition it evolved into a grammatical marker of futurity, and is currently acquiring secondary modal coloring, while the periphrastic future form in (17c), *be going to* $+V$, is taking over as standard. Van der Auwera and Plungian (1998: 111) discuss evidence for the so-called *remodalization cycle* in the history of the future where modal meaning gave rise to the (postmodal) future, which in turn functioned as a (premodal) basis for epistemic necessity as exemplified for the English *will* in (18).

(18) *(doorbell)* That will be the delivery man.

Fleischman (1982) calls this process a bidirectional semantic shift: from modality through grammaticalization to tense, back to modality again (see also Traugott 2006). Next, (17d) and (17e) involve a strong sense of prediction, involving an element of planning. In DS these forms are referred to as futurative progressive and "tenseless" future (after Dowty 1979) respectively. Finally, the overtly modal forms in (17f)–(17i) can also have future-time reference.

(17) f. Lidia must be playing in a concert tomorrow evening.
g. Lidia ought to/should be playing in a concert tomorrow evening.
h. Lidia may be playing in a concert tomorrow evening.
i. Lidia might play in a concert tomorrow evening.

The forms in (17f) and (17g) are epistemic necessity future, which can also be classified as evidential, and those in (17h) and (17i) are epistemic possibility future. In DS, all of the means of expressing the future are assessed with respect to the degree of modality, corresponding to the degree of epistemic detachment from the situation expressed by the sentence. Temporality is represented by means of a version of a sentential operator of epistemic modality, which is a sub-species of the operator of Acceptability ACC, loosely modeled on Grice (2001). ACC assumes the value "epistemic" ($\vdash$) and it is also indexed for the degree of modality represented as Δ as in (19).

(19) $\mathrm{ACC}_{\Delta} \vdash \Sigma$ "it is acceptable to the degree Δ that Σ is true"

ACC is also indexed for a type of expression such as for example *rf* for "regular future" and, where necessary, for the lexical source such as *may* in *epf may* for "epistemic possibility future with *may*", in order to differentiate it from *epf might* or *epf could*. ACC stands for the modal and temporal specification of the situation and in this sense it is a sentential, or a propositional, operator. However, DS has a lot to say about the qualities of the unit it operates on. As is well known, there are problems with making it operate on a sentence or a proposition. Equally, there are problems with making it operate on a state or event (see, for example, Kamp and Reyle 1993; Pianesi and Varzi 2000). DS makes it operate on the so-called merged proposition which is then analyzed as a merger representation Σ and is composed according to the principles of pragmatic compositionality of the merger illustrated above in Figure 9.2. Merged proposition avoids the pitfalls of sentences or propositions in that, founded on the principles which include the rejection of the syntactic constraint, it stands for the representation of the situation by the speaker—or, in more detail, for the main, primary meaning pertaining to the situation intended by the Model Speaker and recovered by the Model Addressee. ACC operates on this content, forming the Σ of the entire utterance, including its temporal orientation.

Figures 9.6–8 are examples of merger representations for future-time reference. Figure 9.6 represents regular future in (17a). In order to differentiate between the representation of the entire utterance and the representation of the situation on which ACC operates, the first is referred to as Σ and the latter as Σ′.

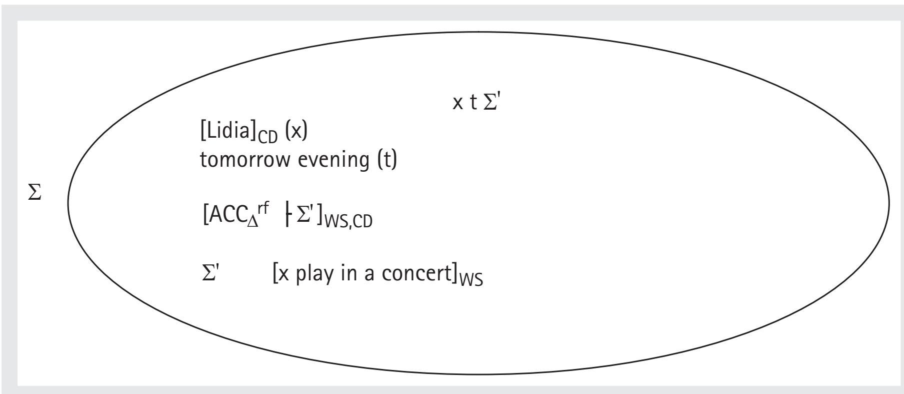


Figure 9.6. Σ for example (17a), regular future

The indices following $[\mathrm{ACC}_{\Delta}{}^{\mathrm{rf}} \vdash \Sigma']$ signal that the relevant, active building blocks of the representation are here word meaning and sentence structure (WS) and cognitive default (CD) in that the meaning of the construction is produced by the meaning of the words, the grammar, and the default value of the auxiliary *will*. WS alone is not sufficient here because of the existence of the wide array of other non-future uses of *will*, to mention only habitual *will*, also called dispositional necessity *will*, as in (20).

(20) Lidia will always play the piano when she is upset.

Futurative progressive in (17d) obtains a merger representation as in Figure 9.7.

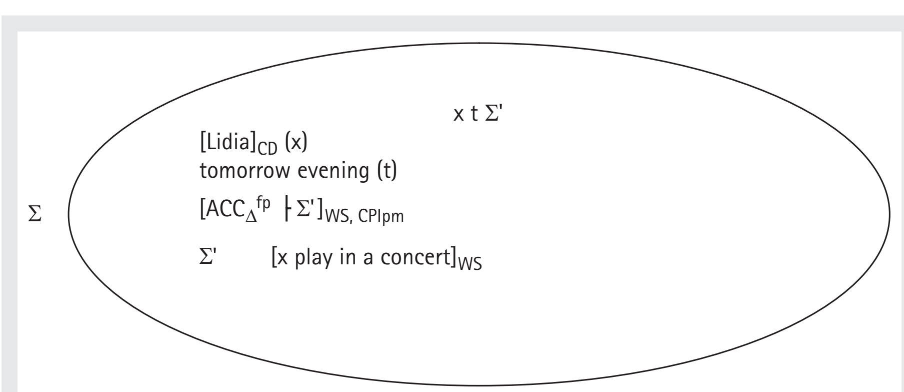


Figure 9.7. Σ for example (17d), futurative progressive

The superscript *fp* on Δ indicates that the degree of epistemic modality pertains to that associated with the futurative progressive (*fp*) form. The subscript CPI_{pm} on ACC indicates here that the future-time reference of the present continuous form "is playing" is obtained via conscious pragmatic inference (CPI) that contributes to the primary meaning (pm).

When future-time reference is represented by means of an appropriately indexed ACC, it is easy to depict the fact that overt modals such as *may* in (17h) also perform future-time reference, amalgamated with the function of conveying modal detachment. (17h) obtains a representation as in Figure 9.8.

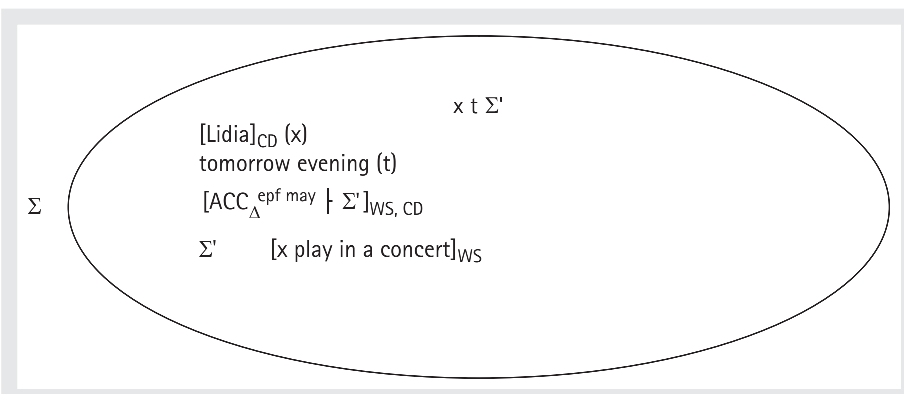


Figure 9.8. Σ **for example (17h), future** ***may***

Epistemic possibility future conveyed by means of the verb *may* (*epf may*) is associated with the cognitive default (CD) type of information and thereby with the source of information IS (properties of the human inferential system) of Figure 9.1. Since *epf* can be conveyed by other verbs, the specification of the lexical source (*may*) is necessary in this case.

It is extensively argued in DS that, just as the future is modal in that it pertains to various degrees of probability and is conveyed as various degrees of speaker's detachment from the situation, so the past is modal in an analogous way.[7] In other words, just as there is branching future in the sense of the mental representation of the future, so there is branching past in the sense of the mental representation of the past, also called the concept of the past, psychological past, internal past, and so forth. The present, although normally associated with the deictic center of a speech event and therefore represented with a high degree of modal commitment, also allows for a cline of epistemic modality. To put it in

[7] See Jaszczolt 2009 for evidence and arguments supporting the view that the semantic (and thereby, in DS, conceptual) category of temporality is supervenient on the semantic category of modality. Merger representations for future, present, and past-time referring constructions are also presented there.

psychological/phenomenological terms, just as the concept of the future pertains to anticipations, so the concept of the past pertains to memories and the concept of the present to the view of what is most likely to be the case now. For example, (21a)–(21e) all refer to the present.

(21) a. Lidia is playing in a concert now.
b. Lidia will be playing in a concert now.
c. Lidia must be playing in a concert now.
d. Lidia may be playing in a concert now.
e. Lidia might be playing in a concert now.

The present continuous form in (21a) yields what is called in DS regular present; (21b) and (21c) exemplify epistemic necessity present, also classified as evidential; the forms in (21d) and (21e) pertain to epistemic possibility present. Dispositional necessity present, also known as habitual present, in (21f), repeated from (20), can also be added to this list.

(21) f. Lidia will always play the piano when she is upset.

Merger representations for (21a)–(21f) are constructed analogously to those for future-time reference, using the ACC operator, the indices on Δ, and the indices for the types of information that builds up Σ. The same principles govern the construction of Σs for past-time reference as in (22a)–(22g).

(22) a. Lidia played in a concert yesterday evening.
b. Lidia was playing in a concert yesterday evening.
c. Lidia would have been playing in a concert then.
d. Lidia must have been playing in a concert yesterday evening.
e. Lidia may have been playing in a concert yesterday evening.
f. Lidia might have been playing in a concert yesterday evening.

The past-time referring expressions "played" and "was playing" in (22a) and (22b) correspond to regular past; (22c) and (22d) are examples of epistemic necessity past, also classified as inferential evidentiality; (22e) and (22f) pertain to epistemic possibility past. We should also add to this list the past of narration in (22g):

(22) g. This is what happened yesterday: Lidia goes to London, meets Sue at King's Cross Station, suggests going to a concert…

The past of narration signals a high degree of epistemic commitment, comparable to that of regular past, in that the situation is vividly present in the speaker's memory. Again, the differences in the degree of modal detachment can be easily captured in DS by means of ACC, degrees of Δ, and sources of information that composes Σ. The analysis is analogous to that presented for the future.

The interaction of WS, CPI, CD, and SCWD, paired with the freedom from the syntactic constraint on the composition of the main meaning of the utterance, give this framework an unquestioned advantage over other post-Gricean contextualist approaches. The task of merger representations of DS is to model the main message that is intended by the speaker and recovered by the addressee, rather than to model an artificial construct of a development of the logical form, variously called what is said, explicature, or a modulated proposition. The latter construct does not correspond to a psychologically real stage in utterance interpretation but rather sits halfway between utterance-type meaning and speaker's primary meaning. In other words, DS takes a big step toward representing psychologically real and psychologically interesting meanings, not shunning abolishing the artificial and totally unnecessary restrictions imposed by the WS source, namely by the logical form of the sentence. Since the logical form of the sentence is not a level of meaning that would correspond to a real phase in utterance processing, the methodological principle of parsimony requires that we treat it on a par with other sources. What happens in practice is that the types of information identified here interact incrementally, without necessarily "waiting" for the logical form of the sentence to be delivered first. Although the full algorithm of this interaction is still a task for the future and we don't at present know the exact length or content of the interacting building blocks, the incremental character of processing and the interaction of the building blocks as identified here are well attested. Temporality is a good example on which this interaction can be demonstrated. It is particularly diaphanous when considered in a contrastive perspective. Languages such as Thai, in which overt marking of temporality is optional, often resort to CPI and CD in conveying the temporal location of a situation (see, for example, Srioutai 2004, 2006).

9.8 Future prospects

There is currently an ongoing debate in post-Gricean circles concerning the extent to which considerations of discourse processing should enter into a theory of meaning (cf., for example, Levinson 2000; Atlas 2006; Jaszczolt 2008). DS stands firmly on the side of psychologism in semantic theory in that it demonstrates that the composition of meaning, understood as the main, primary intended meaning, can only be explained when we take into consideration all of the building blocks that contribute to this meaning, and the building blocks are only explained when we trace them down to the sources of information about meaning that participate in the construction and recovery of a message.

The rejection of the syntactic constraint comes out as a natural concomitant of this pragmatic compositionality view. The cases where merging WS with, say, SCWD or CD produces a representation that does not resemble the logical form of the uttered sentence are no longer regarded as special cases. The preservation of the logical form as the "core" is not the norm, and neither is there a requirement that the primary meaning of the utterance has to entail the proposition uttered or be a development of the logical form of the sentence. Primary meaning is the most salient, main meaning *tout court*. It is rather surprising that a unit as natural as primary meaning was not recognized as an entity in its own right in Gricean pragmatics before DS but instead was split between the explicit/said content and the implicature. If the main intended meaning happens to look like Gricean implicature, so be it: it is still the main meaning to be modeled in a psychologically real account of discourse.

Moreover, artificial efforts to maintain the need for the unit pertaining to the developed logical form can lead to formidable complications of the theoretical apparatus. To give just one example of Recanati's (2004) literal–nonliteral distinction, the preservation of the syntactic constraint leads to the following typology. First, there is *t*-nonliterality (standing for *type*-nonliterality) which pertains to departures from utterance-type meaning, such as, say, meaning that Paul is in Cambridge by "He is here". But this meaning is still *m*-literal (*minimally*-literal) because it departs from the sentence minimally: the departure is constrained by conventions, such as filling in indexical expressions with referents. In contrast, an implicature such as that Paul is easily contactable counts as *m*-nonliteral. Next, Recanati introduces *p*-literality (for *primary* literality), which pertains to departures from the sentence but not necessarily to those governed by linguistic conventions. These departures must be a result of automatic rather than conscious inferential modification. For example, "The Queen visited Cambridge and everyone cheered" has *p*-literal meaning that the Queen visited Cambridge *and as a result* everyone cheered. This meaning is *p*-literal but *m*-nonliteral. Now, to complicate matters further, metaphors are normally *p*-literal in that they involve an automatic shift to the ad hoc, target concepts. "Mary is a doormat" has as *p*-literal meaning that Mary is a subservient woman without strong character or ambitions. But novel metaphors that require conscious processing will count as *p*-literal and at the same time nonliteral in the ordinary sense of being figurative. It appears that they are *p*-literal not in virtue of being automatic, subconscious meanings because the awareness if their "figurative" character and makeup, as Recanati (2004*a*: 77) says, comes in degrees. They are *p*-literal because they have to fit into the previously assumed matrix of developments of the logical form which are not linguistically controlled, i.e., cases that include the discussed enrichment of *and* to *and as a result* or the conceptual shift from *doormat* to *subservient person*. Since novel metaphors which are felt as figurative meanings also pertain to such modifications of the logical form of the sentence, they have to fit into the *p*-literal slot, sense or no sense, confusion or no confusion!

The solution to this conceptual muddle is very simple indeed and comes from DS. Once the syntactic constraint is abandoned, the problem disappears: there is no p-literalness, there is even no m-literalness. Building blocks of the merger representation result in the primary meaning which is *felt* as literal or nonliteral, although the distinction itself is of no interest for composing the representation. On the other hand, the sources and types of information are directly informative and have a commonsense feel to them. It is as simple as that.

The DS-theoretic analysis has now been applied to quite a wide range of constructions spanning several languages. However, the theory is still quite new and in the process of development. There are constructions and languages that have not been discussed. There are also many theoretical questions that require answers, such as the regularities, laws, or heuristics governing the interaction of the building blocks of the merger representation. It seems to me at present that the quest for an algorithm of this interaction can be best pursued empirically; the construction of the theory ends with proposing Σ for various types of expressions and various problematic phenomena such as presupposition. The question of the exact interaction is a question of processing and should be regarded as such: "armchair psychologizing" ends with constructing Σs. The questions of, say, the length of the WS unit which interacts with, say CPI, or the exact list of what counts as CD in virtue of the structure and operations of the brain or as SCWD in virtue of automatization of inferences, are best pursued within experimental psychology. DS provides a semantic theory that allows for processing considerations and it feeds ideas to experimental psychology. It also provides a precise account of what counts as default interpretation, thereby making a big step forward from the terminological misunderstandings between those who take them to be statistically common meanings, automatic additions in processing, shortcuts through inferential efforts but themselves still inferential, commonsense context-dependent interpretations of natural language sentences, and so on and so forth.

CHAPTER 10

DEPENDENCY GRAMMAR AND VALENCY THEORY

VILMOS ÁGEL
KLAUS FISCHER

10.1 Central questions

10.1.1 Contextualization

The best way to introduce a reader to dependency grammar (DG) and valency theory (VT)—both come out of the same research tradition—is to present the syntactic theory of Lucien Tesnière, which is still relevant today (section 10.1). In the following sections we will give an overview of the development, current state, and possible future concerns first of VT (section 10.2), then of DG (section 10.3). However, this chapter can only give a rough outline of DG and VT; a comprehensive research overview can be found in the two volumes of the *Handbook of Dependency and Valency* (Ágel et al. 2003; 2006), which features altogether 121 articles in German or English. The present section is organized as follows: sections 10.1.2 and 10.1.3 introduce Tesnière's DG and VT, while section 10.1.4 outlines the central properties of a modern DG and VT.

The authors would like to thank Tim Pooley for looking through this chapter.

10.1.2 Tesnière's structural syntax

The founder of modern DG and VT is the Frenchman Lucien Tesnière (1893–1954) whose main work *Éléments de syntaxe structurale* was published posthumously (1959; 1976). The *Éléments* had essentially been completed in the 1940s. A short summary under the title of *Esquisse d'une syntaxe structurale* was published in 1953. There is a shortened translation of the 670-page *Éléments* into German (Tesnière 1980), but none into English. Tesnière's theory of structural syntax—only his followers named it "dependency grammar"—has the following basic features:

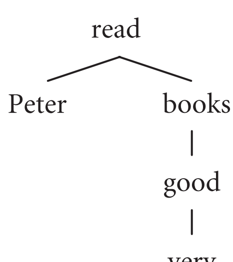
1. It combines a *typological* with a *universal* outlook.
2. It accounts for the *infinity* of theoretically producible and analyzable sentences.
3. It is *autonomous*.
4. It is *functional*.

Tesnière, who spoke more than 20 languages and used examples from more than 60 languages in the *Éléments*, placed great emphasis on applying his descriptive apparatus to different languages and linguistic types. He realized this typological-universal outlook through a modular approach: the *Éléments* divide into three main parts, of which each deals with one of the basic relations of Tesnière's structural syntax: *connexion, jonction, translation*.

The 300-page section on connexions introduces DG and also VT as an organic part of the former. It describes the basic structure of sentences, which is not based on the binary divide into a subject (or NP) and a predicate (or VP) but on the *principle of endocentric verb centricity*. Endocentricity means in this context that the sentence structure unfolds from the lexical verb, that the lexical verb is the precondition for the sentence structure.[1]

Connexions refer to the structurally induced co-occurrence of words, dependency to the hierarchical ordering of connexions. We explain this using the example *Peter read very good books*, whose structure Tesnière would have represented in the following tree diagram (*stemma*):

(1)



[1] Although Tesnière does not use the term *endocentricity*, he not only investigates the respective phenomenon, but his discussion is more consistent than that of Bloomfield (1933: 195), as he assumes that in a sentence such as lat. *filius amat patrem* 'the son loves the father' the predicate and subject are not *amat* and *filius* respectively, but *ama-* and *filius . . . t*. The verb ending *. . . -t* belongs to the verb morphologically only; syntactically it is part of the subject (Tesnière 1976: 104); cf. section 10.2.6.

This stemma is only superficially similar to a constituency tree, it is a *dependency tree*. The branches are not graphical representations of part–whole relations but of connexions: seen from the top, they are government relations, seen from the bottom they are dependency relations. The element at the top, the *regent*,[2] creates a structure for one or several *dependents*. The graphical design of the stemma states that the sentence is primarily about a specific act of reading rather than a specific quality of the person Peter or the object books. It states further that a person as reader, realized as subject, and something that is read, realized as direct object, are involved in the act of reading. The nouns *Peter* and *books*, structurally the direct dependents of the regent *read*, could be involved in different expressions referring to completely different acts or scenes (*Books are exciting, I gave Peter a compact disc player*). The grammatical functions of *Peter* and *books* thus vary dependent on the verbal regent, but not the other way round (verb centrality). The dependency principle can also be applied to the relations between *books* and *good*, and to those between *good* and *very*: *good books* refers to a specific quality of books and not to a bookish quality. By *very good*, we mean the grading of a quality and not a qualified grading. Thus the respective connexions must be such that *books* is the regent of the dependent *good*, and *good* the regent of the dependent *very*.

The constituency structure of the sentence can be derived from the dependency structure: the whole sentence is a verbal node (*nœud verbal*), consisting of the top regent (*read*) and two noun nodes. One of them (*Peter*) just consists of the nominal regent (that is a dependent of the verbal regent at the same time). The other noun node (*very good books*) consists of the regent *books* (that is also a dependent of the verbal regent) and the adjectival node *very good* that consists of the adjectival regent *good* and the adverbial dependent *very*.

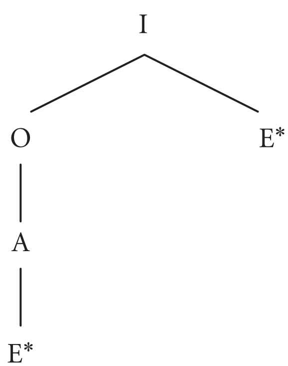
What can—in contrast to modern dependency grammars—only partially be reconstructed, is word order. Tesnière differentiates strictly between the multidimensional structural order of sentences and their one-dimensional linear order, which is not represented in the stemma. However, the translation of linear into structural order (analysis) or of structural into linear order (production) is guided by the linguistic type of a language: Tesnière (1976: 22ff.) differentiates between *centrifugal* languages that linearize the structural order in a descending manner: first the regent, then the dependent (fr. *cheval blanc*), and *centripetal* languages that linearize the structural order in an ascending manner: first the dependent, then the regent (*white horse*).[3]

[2] Note that a *regent* (or *governor*) in dependency grammar governs another word form in the first instance, and only indirectly a phrase. The term is thus used in a slightly different way from *head* in phrase structure grammar. Mel'čuk (2003: 191f.) differentiates between *governor* as the external governing element and *head* as the internal governor.

[3] Tesnière's word order typology anticipates the basic idea of the word order typology of Greenberg (1963).

However, let us return to the structural order. In our example above, all regents and dependents belong to *four word categories*: verb, noun, adjective, and adverb. It is not by accident that it is exactly these four. According to Tesnière, there are universally only the four lexical word classes verb, noun, adjective, and adverb that have node building power. Also universal are the dependency relations that regulate the possible constellations between the lexical classes: the verb is the top regent, it governs but cannot be governed. Nouns are direct dependents of verbs and direct regents of adjectives. Finally, adverbs are direct dependents of verbs (*He read SLOWLY*), adjectives (*VERY good*), or adverbs (*VERY slowly*). This leads to the following *universal dependency structure*:

(2)



Imitating the Esperanto endings, Tesnière introduces the following symbols: I=verb; O=noun; A=adjective; E=adverb. E* stands for a recursive relation, i.e., adverbs can also depend on adverbs. Tesnière calls a stemma that contains symbols a virtual stemma and one that contains word forms an actual stemma.

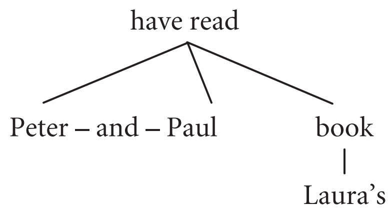
Our introductory example is simple in a literal sense since

(a) all regents and dependents are *simple word forms* (*Peter, read, books, good, very*);
(b) all regents and dependents can be attributed well-defined *grammatical functions*: first noun node=subject=*Peter*, verb=predicate=*read*, second noun node=direct object=*very good books*, adjective node=first grade attribute of first degree=*very good*, adverb node= attribute of second degree=*very*.[4]

But linguistic reality is far more complicated. For instance, predicates often do not just consist of a single verb form but of a verb complex (e.g., *have read*). Nouns can not only be specified by adjectives but also by many other elements (e.g., *Laura's*). And all regents and dependents can be doubled or multiplied through co-ordination (e.g., *Peter and Paul*). The basic structure of sentences can thus be expanded through qualitative and quantitative operations (Tesnière 1976: 80), e.g., *Peter and Paul have read Laura's book*:

[4] Tesnière does not use the traditional functional terms predicate, subject, object, and adverbial but replaces them with the notions of central nucleus, actant, and adjunct according to his basic principle of endocentric verb centricity (see below).

(3)



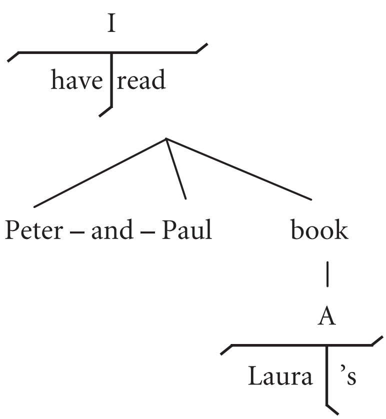
This sentence contains both simple and complex regents and dependents. Regents and dependents have both lexical and structural (grammatical) aspects. To reflect this, Tesnière calls simple and complex regents and dependents *nuclei* (*nucléus*). In the case of a simple nucleus such as *read* in (1) the lexical and grammatical aspects are united in one word form. Dissociated nuclei such as *have read* or *Laura's*, on the other hand, have a structural regent (*have* or *'s*) that contains mainly grammatical information, and a lexical regent (*read* or *Laura*) that contains mainly lexical information. Tesnière's DG thus contains all elements of a modern phrase structure grammar: the concept of a phrase corresponds to the concept of a node. The structural regent of a nucleus is the head, the lexical regent the core of a phrase. If the nucleus consists of just one word form, head and core coincide.

The responsibility for the qualitative expansion of the basic structure by complex nuclei lies with the *basic relation of transfer* (French *translation*), the responsibility for the quantitative expansion of the basic structure by increasing the number of (simple and complex) nuclei lies with the *basic relation of junction* (French *jonction*). While connexion (dependency) is the inter-nuclear relation of super- and subordination, *junction* is the inter-nuclear relation of co-ordination. Junction is a quantitative phenomenon that can be compared to addition or multiplication (Tesnière 1976: 324). It operates either with a junctive such as *and* (*Peter and Paul*) or without junctive (*veni, vidi, vici*).

Much more complicated is *transfer*. As there are only four categorial types of nuclei, whose possible dependency relations are strictly regulated, all types of complex structures have to be converted into these four categorial types. Transfer is an intra-nuclear relation that consists of two types of operations (Tesnière 1976: 364): (1) change of category, and—as a consequence of this—(2) change of function.

Our examples for transfer are the complex nuclei *Laura's* and *have read*. As the nucleus *Laura's* in the noun node *Laura's book* occurs as a dependent of the noun *book*, according to the universal dependency structure it has to belong to the category of adjective. Likewise, the nucleus *have read* must belong to the category of verb. But now consider that neither can *Laura's* belong to two word classes (noun and adjective) at the same time nor can a complex nucleus comprising at least two words belong to a single word class! Therefore the categories of *Laura's* and *have read* cannot be the genuine word classes of adjective and verb respectively but must be categories that have been created as a result of an operation. The theory of transfer describes the mechanisms and types of such category creation processes. Each transfer features a transferent (French *transférende*), a word the category of which is to be changed, and a translator or translating agent (French *translatif*) which does the grammatical work of effecting the category change. The latter can be a free, bound, or zero morpheme. The result is the translatum (French *transféré*), the operatively created nucleus: the clitic *'s* changes the noun *Laura* into the adjectival nucleus *Laura's* (O > A), which as a result of the transfer is structurally dependent on and functionally attributed to the noun *book*. In the case of the translatum *have read* the verbal transferent *read* has—with the help of the translator in the form of the auxiliary *have*—been transferred from a verbal subcategory, in modern parlance a synthetic verb form, into another verbal subcategory, an analytic verb form; thus I > I (verb is transferred to verb).[5]

(3a)



At the end of this section we would like to come back to the four basic features of Tesnière's DG. We hope to have shown in our sketch that Tesnière's DG is a

1. *typological-universal* theory
2. that is designed to account for the *infinity* of theoretically producible and analyzable sentences.

But what does the claim mean that Tesnière's structural syntax is

3. *autonomous* and
4. *functional*?

By autonomy of his syntax Tesnière means that syntax, which investigates the inner form of the sentence, is conceptually different from morphology, which investigates the outer form of the sentence (Tesnière 1976: 34). Both should thus be kept separate in description: syntax adheres to its own laws.

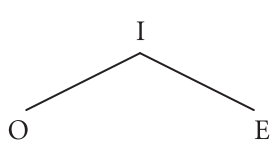
[5] Tesnière chooses a T to represent transfer graphically. The translatum (or its symbol) is placed above the horizontal T-bar, transferent and translator are placed below the horizontal T-bar to the left and right of the vertical T-stroke (or the other way round).

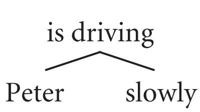
Syntax is autonomous since it is as clearly differentiated from semantics as it is from morphology (Tesnière 1976: 40ff.). Here Tesnière anticipates the argument of the early Chomsky: he sees the best proof for the independence of syntax from semantics in semantically absurd sentences such as *Vertebral silence upsets the lawful sail* (*le silence vertébral indispose la voile licite*), which are structurally correct, possessing the same structure as meaningful sentences.

The autonomy of syntax from semantics does not mean for Tesnière that semantics is linguistically of no interest. On the contrary: meaning is the ultimate *raison d'être* of structure and thus the *indirect subject* of syntax (Tesnière 1976: 40). We have learned from connexion theory and its interplay with transfer theory that Tesnière's syntax theory is functional. The simple nuclei in the introductory example (1) allowed a direct functional interpretation of the nuclei and nodes. To ensure that a functional interpretation of complex nuclei such as (3) is guaranteed, nucleus categories created by transfer were permitted in addition to the original nucleus categories. In Tesnière's thinking, structure only exists to the extent that there are functions. Thus his structural syntax is also a functional syntax (Tesnière 1976: 39). We interpret Tesnière's functional outlook as a logical and consistent attempt to motivate the autonomous grammatical structure by establishing a theoretical link between semantics and grammatical functions.

10.1.3 Tesnière's valency theory

In his famous *drama metaphor* Tesnière (1976: 102) compares the sentence to a small drama. In a similar way to a drama encompassing an action or event, actors and (temporal, spatial, etc.) circumstances, the verbal node contains a central nucleus, which, according to the universal dependency structure, can govern two types of nodes: noun nodes and adverb nodes:

(4)


(5)


The central nucleus corresponds to the drama event: it is a (simple or transferred) verb (*is driving*), which in modern VT is called a *verbal valency carrier*. The functional equivalent of the drama actors are the *actants*, which are noun nodes (*Peter*). In modern VT *actants* are also called *complements*. The functional equivalent of the circumstances are the *circumstantials* (French *circonstants*), which are adverb nodes (*slowly*). In modern VT they are also called *adjuncts*.

The traditional principle of the binary divide of the sentence into subject (or NP) and predicate (or VP) is thus replaced by the principle of *verb centricity* that establishes a *functional threepartite division of the sentence* into verbal valency carrier, actant, and adjunct. This functional division is secured in the categorial division into verb (equivalent), noun (equivalent), and adverb (equivalent). However, actors and circumstances, or actants and adjuncts, do not have an equal input into a drama or a sentence respectively. The constitutive elements of a drama are event and actors, the constitutive elements of a sentence the central nucleus and the actants. The circumstances in a drama or the adjuncts in a sentence, on the other hand, have the external function of framing the constitutive participants. There is only a dependency relation between I and E, while between I and O there is a *valency relation* in addition to the dependency relation. Tesnière uses a famous comparison, his *atom metaphor*, to introduce this additional relation. According to this metaphor one can compare the verb to an "atom with a particular number of hooks that can—according to the number of hooks—attract a varying number of actants, which it keeps in its dependence. The number of hooks that a verb possesses, and consequently the number of actants that it governs, constitutes what we call the valency of a verb" (Tesnière 1976: 238).[6]

Valency is thus according to Tesnière the *number* of the *potential* actants of a verbal valency carrier. The actants are anchored in verb meaning. For instance, French *donner* or English *to give* demand three actants. These verbs are trivalent, notwithstanding whether apart from sentences such as *Alfred donne le livre à Charles* (Tesnière 1976: 107; 'Alfred gives the book to Charles') there are also sentences such as *Alfred donne aux pauvres* ('Alfred gives to the poor') or *Alfred donne la main* ('Alfred holds hands') that only realize part of the valency potential of *donner* (Tesnière 1976: 239).

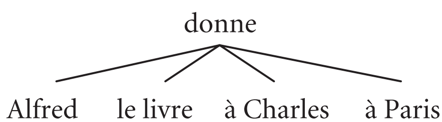
Tesnière distinguishes between three kinds of actants (Tesnière 1976: 107ff.). All three have the *dependency grammatical form* in common: they are noun nodes. They also share *proposition formation* with the verbal nucleus. Their difference lies in their *semantic relation* to the verbal nucleus: the first actant, the traditional subject, is the actant that carries out an activity (in the active clause), the second actant is the actant to which an activity or action happens, and the third actant is the one to whose benefit or detriment something happens. The later concepts of deep cases, semantic or theta roles are easily recognizable in this description.

Tesnière, who, in keeping with his typological interests, runs the different formal types of actant realization of various languages past the reader, does not overlook the problem of *complement–adjunct differentiation* that has come to occupy

[6] On peut ainsi comparer le verbe à une sorte d'**atome crochu** susceptible d'exercer son attraction sur un nombre plus ou moins élevé d'actants, selon qu'il comporte un nombre plus ou moins élevé de crochets pour les maintenir dans sa dépendance. Le nombre de crochets que présente un verbe et par conséquent le nombre d'actants qu'il est susceptible de régir, constitue ce que nous appellerons la **valence** du verbe. [Bold in the original]

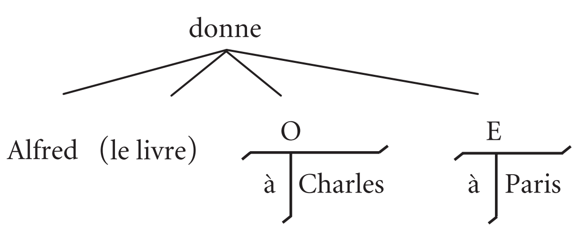
such a central place in modern valency theory. Particularly problematic here are the prepositions. According to the universal dependency structure, they have to be translators for both nominal and adverbial translata. For example, in the sentence *Alfred donne le livre à Charles à Paris* the chain *à Charles* is a noun equivalent, while the chain *à Paris* is an adverb equivalent:

(6)



Here, valency criteria that are independent from the valency structure must decide why *à Charles* constitutes a transfer of O (*Charles*) > O (*à Charles*), while *à Paris* one of O (*Paris*) > E (*à Paris*):

(6a)



Tesnière can draw here on semantic criteria (proposition formation, semantic role) and on the fact that adjuncts are optional on principle. Also, the chain *à Paris* can be replaced by a single adverb (*ici, là*), but not the chain *à Charles*. Tesnière uses an additional typological argument: in case languages the chain *à Charles* corresponds to a single noun in the dative (Tesnière 1976: 114f.).

Finally, we would like to address the question of Tesnière's relevance for modern VT and syntax:

1. Tesnière does not only demonstrate that a syntactic theory without VT is impossible, he also derives his VT from his syntactic theory in a logical and consistent fashion. Conversely he shows that valency provides an additional motivation for the universal dependency structure in that the valency relation operates on the government relation of I to O.
2. In Tesnière's work, all fundamental questions of modern VTs are addressed extensively and instructively: (a) the problem of the valency carrier; (b) the problem of complement–adjunct differentiation; (c) the distinction between valency and its obligatory or optional realization; (d) the problem of structural valency realization, i.e., the question to what extent the structures of individual languages influence the forms and types of possible actants.
3. As languages like English or Chinese mark actants not just morphologically or prepositionally but also positionally, the problem of structural valency realization concerns the fundamental question of DG of whether Tesnière's rigid separation between structural and linear order should be maintained. Among the

modern dependency grammars, there are projective grammars which preserve word order (cf. Eroms and Heringer 2003).

10.1.4 Fundamental traits of dependency grammar and valency theory

All grammatical theories assume that multidimensional structures are hidden behind the temporal sequences of spoken signs. Their architecture cannot be simply read off the spoken data, nor does a surface-oriented morphological description capture it: a syntactic theory must thus provide own categories and relations to describe this architecture. Tesnière's *autonomy principle* is based on the *categorial independence of the syntactic theory component*. His dependency grammar reflects semantic (and pragmatic) differences to the extent that they can be mapped onto syntactic categories, but syntax is not reduced to the expression of semantics (and pragmatics). This non-reduceability of syntax follows not only from the impossibility of projecting every semantic (and pragmatic) difference and opposition onto syntactic categories but also from the mechanisms of language change and grammaticalization. DG and especially VT put great emphasis not just on the *typological* but also on the *linguistic-historical adequacy* of theory formation. The theory should not just be applicable to a few modern literary languages but to any language at any historical stage.

The linguistic forms *underdetermine* the sentence structure. This is one of the reasons why this handbook presents different grammatical theories, and not concepts of a unified theory of grammar. A second reason is that there is no consensus on which *concepts* are the basic ones for describing linguistic structure (e.g., dependency or constituency) or even which concepts are necessary (e.g., functional notions such as subject and/or formal ones such as NP). Another question is how complex or abstract should the structural descriptions be. The crucial point is that a methodologically controlled link between theory and empirical data must be guaranteed. This challenge is closely connected to the description and explanation of grammatical constructions: How should a syntactic theory deal with these? Should it take (a) a projectionist perspective (from the elements to the whole), or (b) a constructionist perspective (the other way round)?

DG is fundamentally a projectionist theory, which describes structure starting from the individual lexemes: DG is a word grammar. Individual lexemes have a combinational potential or tolerance, which creates or does not prevent or excludes structures. DG is thus, in this respect like HPSG, not a "slot theory" (Feilke 2004: 54), in which the lexicon merely provides fillers for lexically independent syntactic structures. An essential part of grammar is in the lexicon: in the potential of lexemes for connexion, junction, transfer, and valency. This word grammar perspective of DG has a number of consequences:

1. The combinational potential of individual lexemes is founded on their meaning and their syntactic potential. But how does the autonomy principle of DG interact with this lexical foundation? In most cases the dependency branches can be directly semantically interpreted. But no identity or one-to-one correspondence between the structural and the semantic level is assumed, as for instance in categorial grammar and in the basic outlook of functional approaches. As a structural theory, DG assumes an independent structural level. Only with this assumption is it possible to ascertain the degree of correspondence between syntactic and semantic structures (Welke 1988; Helbig 1992). The word-grammatical perspective, on the other hand, ensures that the degree of syntactic independence is as small as possible, and that syntactic structures are motivated semantically. DG shares this search for semantic motivation with categorial grammar and functional theories, in spite of the differences outlined above.
2. DG's lexical perspective, i.e., its taking individual lexemes as the point of departure, is compatible with its interpretation as either a reception or a production grammar. The view of DG as a reception grammar is possible because a recipient's knowledge of the syntactic–semantic potential of individual lexemes, supported by framing and priming, leads to reliable analyses (Heringer 1984; 1985; Gansel 2003). The view of DG as a production grammar is possible because the production of utterances depends on the choice of individual lexemes both on the basis of communicative intention and—trivially but crucially—on the basis of the syntactic–semantic opportunities and restrictions that have been conventionalized in the vocabulary of an individual language. For instance, the English verb *lie* can hardly be combined with a *that*-clause expressing the content of a lie (Fischer 1997: 241), while a corresponding construction is acceptable in Hungarian:[7]

    (7) English ?He lied that he was poor.

(8)	Hungarian	Azt	hazudta,	hogy	szegény.
		$\text{that}_{\text{Pron.Acc}}$	lied-he	$\text{that}_{\text{Conj}}$	poor

3. DG is a data-friendly syntactic theory in the sense that it is not restricted to the representation of a few abstract structures but aims at the description of a considerable sample of natural language. DG is also a data-friendly syntactic theory in the sense that it aims at a level of abstraction that still permits a methodologically controlled link to the data.
4. The word-grammatical perspective makes DG particularly suitable for natural language processing (parsing, generation, translation).
5. The relative simplicity of dependency grammar representations makes DG suitable for didactic purposes. DG representations are used in a number of learner grammars and German as a foreign language material.

[7] Acc = accusative, Conj = conjunction, Pron = pronoun.

From the lexical perspective it follows that the central question of DG is the justification of dependencies, i.e., the motivation of connexions between words and their hierarchical ordering. In modern DG both are—to a greater extent than by Tesnière—justified using valency relations (e.g., Engel 1994). Nevertheless, both are not identical: DG is a general syntactic theory which, according to Welke (1995), can be founded on three types of dependency relations: government, endocentricity, and subcategorization (cf. 10.3.3). VT is a partial theory, which is concerned with the combinational potential and realization of relational signs (verbs, nouns, and adjectives). VT cannot but be integrated into other syntactic theories, its integration is essential: no grammar with comprehensive descriptive and explanatory ambition can do without a valency component.

We conclude this section with the provocative question: "Is DG the best of the theories presented in this handbook?" In spite of the many advantages that DG has over its competitors and that are demonstrated in this chapter, a positive answer would be premature. Neither DG nor other theories can presently lay claim to the title of *best theory*. As Jacobs (2008) shows convincingly, there are a number of central, i.e., *not* peripheral, grammatical phenomena that can only be described adequately from a constructionist perspective. A (different) number of equally central grammatical phenomena, on the other hand, require a projectionist perspective. A convergence of projectionist and constructionist grammatical models is definitely required (cf. 10.2.5, 10.2.7, and 10.3.4). The future of DG—as that of other theories—thus depends in great measure on whether the DG framework (or that of another theory) allows the development of a theoretical format that achieves the integration of projectionist and constructionist components—not just as ad hoc stipulation but in an empirically and methodologically sound way.

10.2 Valency theory

10.2.1 Contextualization

Human language does not construct its messages holistically but in predicate–argument structures: events are perceived as consisting of things and relationships between these things. Nouns refer to entities, among which a verb as a relational expression creates a link. The referents of verbs are more abstract and can only be perceived indirectly: for instance, you can only point at them by pointing at one or several objects (cf. Leiss 2002).

All languages possess linguistic means to create predicate–argument structures (proposition formation). Valency theory (VT) investigates this essential core of human language, both as a universal and as a particular shaped by the characteristics of individual languages.

Modern VT derives from Tesnière, but the valency idea is much older. K. Bühler (1982 [1934]) and J. W. Meiner (1971 [1781]), for instance, are seen as direct precursors, but valency-like concepts have been identified in antiquity, in the Arabic grammatical tradition and among the medieval modists. Both Tesnière and his followers see themselves in the tradition of European structuralism, but the fact that the development of post-Tesnière VT was initially centered on central and eastern Europe is not just a historical accident but can be explained by the obvious application of VT to case languages. By now there exists valency research on a multitude of languages, especially on Romance languages and on English. The number of publications on valency is estimated to comprise around 3,000 titles.

The development of VT did not take place in isolation from the predominantly Anglo-Saxon linguistic approaches that dominated the linguistic scene in the second half of the twentieth century: the semantic feature analysis by J. Katz and J. Fodor (1963), the deep case (semantic role), later scene concepts of C. Fillmore (1968; 1977*a*), the general turn to semantic-pragmatic approaches in the 1970s (generative semantics, functional grammars) should be mentioned as well as categorial grammar, the typological opening up of VT in the last decade, and the recent discussion of concepts from construction grammar (Willems and Coene 2006). Valency theoreticians see related concerns realized in (other) word grammars (Hudson 2006; Starosta 1988) and in functional approaches (Halliday 1994; Dik 1978; Givón 1995).

VT's openness to concepts developed outside its framework is hardly reciprocated by other approaches (but cf. Huddleston and Pullum 2002). However, grammars today are highly likely to possess a valency-type component (of variable analytical depth), usually under a different name (e.g., subcategorization). These valency-type components have moved closer to VT. This is particularly obvious in the X-bar template of generative grammar. Recent modeling of first language learning, which emphasizes the lexeme-specificity of the initial acquisition of syntactic functions (Verb Island hypothesis, Tomasello 2003: 117), supports the word grammar outlook of VT.

Four basic questions of VT can be identified (Ágel 2000: 115ff.):

1. What counts as a valency carrier?
2. What is the complementation potential of a valency carrier?
3. Which forms and types of valency realization exist in different languages?
4. How is valency realized in texts?

In section 10.2.2 we will discuss extensions of the valency concept after Tesnière, followed by discussion of questions 1 and 2 in sections 10.2.3–10.2.5. Section 10.2.6 deals with the typological question 3. For reasons of space, question 4 will only be marginally touched upon. And in sections 10.2.7 and 10.2.8 we will briefly address valency change and applied valency research.

10.2.2 Extensions of the valency concept after Tesnière

10.2.2.1 *Fundamentals*

VT is based on the very simple idea or observation that words pre-determine their syntactic and semantic environment:

(9) The man is washing his shirt.

The verb *wash* creates two slots that require a particular categorial filling in a particular function. The verb allocates the function of a washer (agent) to the first slot and the function of a washed (patient) to the second slot: the washing scene is thus seen from the perspective of the washer, the verb *wash* opens a perspective on the washing scene (Welke 2005).

The functional requirement of the verb largely determines the categorial filling of the slots. In literal usage, only entities with the feature "material" can wash or be washed. And only an entity that can carry out an activity can wash. Correspondingly, the categorial filling of the slots is restricted: subject and object must be realized as NP (noun nodes), not, for instance, as PP (prepositional nodes).

The semantic restrictions here are a direct reflex of our knowledge of washing events. But lexemes are arbitrary in the sense that reality does not force us to form a particular linguistic notion. Reference to washing events would have been possible using a generic verb and an adverbial phrase (*clean with water*). Also, the exact delimitation of what we can refer to using the verb *wash* is not prescribed by reality. In English, cleaning the floor with water may be referred to as *wash(ing) the floor*, but the German equivalent of *wash* is usually not used in this context, but the verb *wischen* ('wipe') is. If one also takes metaphorical usage into consideration, English *wash* is used less widely than German *waschen* (*Geld waschen* vs. *launder money*, cf. Feilke 1998: 75).

Verbs are not just combined with complements but also with adjuncts:

(10) The man is washing his shirt *in the kitchen*.

The PP *in the kitchen* situates the washing event. We know that all physical events take place at a certain place and time, but this is world knowledge and not a requirement of the lexeme *wash*. Sentence (9) above represents a complete linguistic realization of the scenario[8] created by the verb *wash*: no local or temporal information is assumed or must be supplied. This is different in

[8] The term *scenario* refers to linguistically created event types or events, not to language-independent states of affair in reality.

(11) He put the shirt *in the kitchen*.

(12) *He put the shirt.

In (11) the adverbial of location (*in the kitchen*) is constitutive for the *putting*-scenario: *put* opens a slot for an adverbial of location. No minimal *putting*-scenario can be realized without this slot being filled (cf. 12).

We have now prepared the ground for offering a valency description of the combinatorial potential of the verb *wash*:

(13) *wash* <AGENT: material (living being, institution, machine) → *subject*: NP;
PATIENT: material → *direct object*: NP>

VT emphasizes that the different syntactic functions such as subject, direct object, obligatory adverbial, etc. have something in common: to fill verb slots. Therefore, including the subject, they are subsumed under the notion of *complement*, though they are very different in other respects.

10.2.2.2 *Formal marking of valency*

Tesnière restricted valency to the number of complements. In the understanding of modern VT, verbs determine both the number and *kind* of slots (e.g., Helbig 1992). The most tangible expression of valency exists perhaps in the marking of complements using case:[9]

(14) German Der$_{\text{Nom}}$ Hund folgt dem$_{\text{Dat}}$ Mann.
'The dog follows the man.'

In German the follower must be realized in the nominative, the followed in the dative, with cases mainly marked on the determiner. Permutation of the unmarked order "subject before dative object" changes the information structure, but not the event expressed:

(14a) Dem Mann folgt der Hund.
'The dog follows the man.'

Positional marking of nominal complements often occurs in languages lacking case. For instance, the unmarked order of nominal complements in English was grammaticalized in parallel to the loss of case:

(15) The dog follows the man.

(16) The man follows the dog.

The follower occupies the preverbal, the followed the postverbal position. Contrary to German, an exchange of positions (16) changes the stated event.

An additional formal marker is provided by verb-specific prepositions, which are often considered as case equivalents:

9 Nom = nominative, Acc = accusative, Gen = genitive, Dat = dative.

(17) She insists on/*under/*in/*for treating everybody with courtesy.

Two important questions arise from the formal marking of complements:

1. To what extent is formal marking unequivocal?
2. Does formal marking just differentiate between complements, or does it also carry semantic information?

Neither position nor morphology offer unequivocal marking: although positions as such are unambiguously defined (before X, after X), they are not restricted to marking just one function. The definition of English nominal complements therefore has to be based on unmarked or canonical positions (cf. Huddleston and Pullum 2002: 225f.). Morphological marking systems have a tendency toward form syncretism (often subject–direct object; cf. Plank 1984). On the other hand, the number of differentiations that can be achieved in case systems is potentially less restricted than the number of grammaticalized positions.

It would be uneconomical if formal marking were merely to serve complement differentiation, with the mapping of formal markers onto semantic roles happening in an idiosyncratic fashion, e.g., an agent being marked unsystematically by a variety of positions, cases or prepositions. In fact, the mapping follows regularities that are both based on competing principles and are sensitive to the presence of other complements.

10.2.2.3 *Adjectival and nominal valency*

Not only verbs have valency but also adjectives (*proud **of his children***, *keen **on debate***) and nouns (*protection **from disease, Peter's** sister*). Individual adjectives need additional information to express their quality, e.g., of whom or what somebody is proud. Likewise, the constitution of entities sometimes requires further protagonists, e.g., protection *from something*. Although adjectival and nominal valency is frequently deverbal (*bore sbdy → boring for sbdy*; *protect sbdy from sthg → protection of sbdy from sthg*), this is not always the case, as the initial examples show. There are currently three competing approaches in existence in the theory of nominal valency. Only one of them, the so-called nominalization approach, assumes that there is no genuine nominal valency but that the nominal valency structure is inherited from the verb. This approach, which is comparable to X-bar theory, is seen as refuted in valency theory (Teubert 2003).

10.2.3 Valency potential I: Complements and adjuncts

10.2.3.1 *Valency as a unified phenomenon*

For a long time, it had been assumed in VT that valency was a unified phenomenon. Therefore, the delimitation of complements from adjuncts (c/a-delimitation) was seen as fundamental to the development of VT. Considerable time was invested into attempts to operationalize c/a-delimitation (which seems intuitively very clear), but no consensus was achieved. The problem of c/a-delimitation derives from a number of phenomena:

1. Complements can be optional, thus optionality is not a sufficient criterion for adjunct status:

(18) She is reading a book.
(19) She is reading.

The direct object *a book* is "truly optional" (indefinite deletion) since the patient of the reading process does not have to be contextually given for (19) to be understood.

2. As a phrase can function both as a complement and as an adjunct, there is no straightforward formal criterion for c/a-delimitation. The same chain can be c or a, depending on the verb (see *in the kitchen* in (10) and (11) above) or depending on the scenario that an utterance is intended or interpreted to realize:

(20) She meditated *yesterday evening*.

(20) is syntactically ambiguous, since *yesterday evening* can be interpreted as c (direct object) or as a (temporal adverbial).

Solutions to the puzzle of c/a-delimitation (with respective tests) were attempted on a morphosyntactic (complements are formally determined by the verb; Eisenberg 2006: 33ff.), distributional (complements are specific to subclasses of verbs; Engel 1988), and semantic basis (verb and complements constitute a minimal proposition; Fischer 1999), but none proved wholly satisfactory. Nor did the construal of the c/a-transition as gradual or scalar (e.g., Heringer 1984 and 1985; Somers 1987) seem to adequately reflect the nature of the c/a-difference.

10.2.3.2 *Multidimensional valency models*

This research situation was given a surprising twist by Jacobs (1994). He claimed that there were seven (later five) valid operationalizations of the valency concept, to which mutually independent relations corresponded: valency was not a unified phenomenon at all but a cover term for these individual *valency relations*, which, though independent, prototypically occur together. Jacobs's intervention resulted in multidimensional (or modular) valency models, which provide the dominating paradigm of VT today (see also Jacobs 2003). The following outline of a multidimensional model is based on Zifonun et al. (1997: 1,026ff.), but adds the relation "position", thus listing the valency relations needed for the description of English and German. The relations can only be presented but not discussed in any detail:

a) Form relations

(i)	Obligatoriness:	*He devours* ***a hotdog***. **He devours.*
(ii)	Form determination:	*They can see* ***him***/****he***.
(iii)	Constancy:	*She insists* ***on***/****in***/****under***/****for meeting us***.
(iv)	Case transfer:	German *Sie besteht auf* ***einem***$_{\text{Dat}}$/****ein***$_{\text{Acc}}$ *Treffen*. 'She insists on a meeting.' German *Sie freut sich auf* ***ein***$_{\text{Acc}}$/****einem***$_{\text{Dat}}$ *Treffen*. 'She is looking forward to a meeting'.
(v)	Position:	*He devours a hotdog*. **A hotdog devours he.*

The verb determines whether and under which conditions a complement can be omitted (i), and which form features it has, i.e., in which case it has to occur (here: the DO must occur in the objective case) (ii), which preposition provides the link to the object (iii), and, for constant or verb-specific prepositions that can govern different cases, which case is chosen (iv): the German preposition *auf* governs either the accusative or the dative. The choice in (iv) depends on the respective governing verb. Position (v) plays a greater role in a language such as English (cf. 10.2.2.2) than in languages with flexible complement order (e.g., Russian). The English positions are allocated by the verb as a category (subject) or by the subcategories of transitive (direct object) and ditransitive verb (indirect and direct object). Accordingly, English adjuncts possess greater positional variability (see Huddleston and Pullum 2002: 225f.).

b) Meaning relations

(vi)	Proposition formation:	***They*** *arrive* ***at home***.
(vii)	Perspective:	*The car costs* ***$3,000***. Cf. *She buys a car* ***for $3,000***.
(viii)	Synsemantic coding:	*He spies* ***on her***. Cf. *The fly is* ***on her***.

To create an arrival event, it is necessary to refer to an entity that arrives and to a place where it arrives (vi). The verb *cost* puts the price more strongly "into perspective" than the verb *buy* does (vii). The semantic relation between the verb *spy* and the object of *spy* is not created by the spatial meaning of *on* but depends on the verb (viii). The relation between the fly and the person, however, is captured by the meaning of *on*.

The more relations can be attributed to a phrase, the stronger is its claim to complement status. Prototypical complements and adjuncts show convergence of the form and meaning relations. There are implications between the relations. For instance, proposition formation implies that the respective phrase is strongly or weakly in perspective.

All the relations show variability across languages. As we have shown above, the form relation position is important in English, while the role of form determination has been much diminished and case transfer has ceased to play a role at all. These relations are not equally important in determining the c/a-status of a phrase. Proposition formation, if suitably defined and not restricted to arguments (cf. 10.2.3.3), is at the core of valency as a universal phenomenon and decides the c/a-status of phrases (nodes) in most cases (cf. Fischer 1999). Form relations concern the *realization* of valency in individual languages (cf. 10.2.6). The other two meaning relations (perspective, synsemantic coding) are a function of the form relations, which in their turn are—at least up to a point—informed by the meaning of the verb.

10.2.3.3 *Complements and arguments*

Although valency is grounded in predicate–argument structures (and thus is fundamentally a semantic-pragmatic phenomenon), it cannot be equated with these structures. Not all arguments are realized as complements, and not all complements are arguments. Adjectives, for instance, have an argument slot that is not filled by a complement inside the adjective phrase but by the noun to which the adjective refers (*a proud* ***girl***, ***she*** *is proud of her children*). Semantic valency slots can also be realized as part of lexical meaning or be blocked from realization: verbs such as *hammer*, *nail* have incorporated the instrument into the verb. A number of German prefix verbs prevent realization of the patient (e.g., *zuschlagen*: *Er schlägt zu*. 'He hits at somebody or something' vs. **Er schlägt* ***ihn*** *zu*. 'He hits him'), although a patient is assumed.

VT thus cannot be reduced to the predicate calculus, though both take the propositions of linguistic utterances as their point of departure.

10.2.4 Valency potential II: The valency carrier

10.2.4.1 *Identification of the valency carrier*

The determination of valency (or the valency relations) of X assumes that we have identified X (the valency carrier):

(21) The woman is washing her blouse.

What is the relational sign in (21)? The system element *wash* or an instance (token) of the word *wash* in a text (e.g., *washing*) or a verb form of *wash* (e.g., *is washing*)? Do we assume just *one* valency of *wash* or different valencies of different partial paradigms (e.g., active, passive) or even of individual word forms? Does a polysemous verb such as *realize* (cf. 10.2.4.2) have one valency carrier or several? Then there is the question of how to incorporate idiomatic expressions into VT. Is the valency carrier of

(22) He kicked the (*big) bucket/*frying pan. ('died')

the unanalyzed monovalent valency carrier *kick the bucket*? Or should an "internal valency", derived from the verb *kick*, be assumed inside the idiomatic expression?

Realistically, idiomaticity cannot be excluded from VT: not only is the free combinability of elements rather the exception but also idiomatic expressions do possess limited variability (see Ágel 2004):

(23) She kicked up a (big/small) fuss/row/stir/stink/?an argument/* a problem.

Further delimitation problems arise from prepositional and particle verbs: does *after* in *look after somebody* belong to the valency carrier or to the complement? (Does the verb *look after* govern a direct object or the verb *look* a prepositional object? Cf. Quirk et al. 1985, Emons 1974). Is *on* in *put on a dress* part of the valency carrier or an adverbial complement?

10.2.4.2 *The problem of readings*

In addition to the *external* identity problem, the delimitation of the valency carrier of verbal idioms and particle verbs, there exists an *internal* identity problem. For most verbs the simple equation of "one valency carrier = one verb" does not work as verbs tend to have different readings (cf. the *Variantenproblem* in Ágel 2000):

(24) She realized her plan. ['make real']

(25) She realized that her purse was missing. ['notice', 'become aware of a reality']

Although the readings share a meaning that may be identified as "real" (if suitable paraphrases of the readings are chosen), it is not possible to derive from this meaning element and the function of the word formation suffix *-ize* that the verb *realize* can both be used as an action verb with an agent and patient (24) and, more frequently in current usage, as a verb of perception with an experiencer and a stimulus (25). Accordingly, it will hardly be possible to subsume both readings under one paraphrase. Nevertheless, the two readings are not just homonyms but readings of one verb that are semantically connected.

Given our findings, how can it be maintained that the lexeme *realize* determines number and kind of its complements? The lexeme does not seem to achieve this. At least the kind of complements can only be derived if the *meanings of the individual readings* serve as the starting point. If one starts from the lexeme *realize*, it is the direct object that determines the verb: if it refers to a fact (e.g. "that her purse was missing" in 25), the reading "notice" is chosen.

But the problem of readings reaches even deeper. How many readings does the verb *realize* have? The *Longman Dictionary of Contemporary English* indicates five, *Collins Dictionary of the English Language* seven, without there being a straightforward mapping between the two sets of readings.

We conclude that verbs do not only determine their environment but that they are also determined by it: the meaning of verbs is often abstract as verbs have a number of usages that are interconnected through family resemblances. Thus their interpretation needs input from both the linguistic context and the situation. This explains the strong variability of verbs across languages, i.e., cross-linguistic matches are between verb readings rather than verbs, and the (unmarked) sentence-final position in almost half of the world's languages. This preference for positioning the verb *after* the complements can only be reconciled with the VT claim to verb centrality, i.e., the view that the verb *determines* the construction, if the problem of readings is considered. The relationship between verb and complements thus should be seen as a dialectic or two-stage process, in particular in reception: 1. determination of the verb (choice of verb reading); 2. determination of the complements by a verb reading.

The relationship between inherent lexical meaning, combinational meaning, and readings is currently very much discussed in VT, also with reference to construction grammar (see Willems and Coene 2006; Coene 2006).

10.2.5 Valency modification and alteration

10.2.5.1 *Basic valency and valency modification*

For a verb such as *wash*, it can be convincingly demonstrated that its lexical meaning determines its valency. In 10.2.2.1 we have shown that a divalent realization derives from the meaning of *wash*. We consider this realization as the *basic valency* of wash (cf. Welke 1988). From here, a number of systematic *valency modifications* (valency reductions and increases) can be expected: medial usage that is often not formally marked in English (*She is washing* 'washing herself'), secondary subjectivization of the patient (*These shirts wash well*), resultative usage (*He is washing his clothes to shreds*), instrumental usage (*She is washing the car with a hose pipe*), benefactive usage (*She is washing the car for her*) and, in suitable contexts, the indefinite usage (*He is washing* 'doing the washing'). We cannot demonstrate this in detail here, but both the valency potential and the valency realizations of *wash* can be derived from its inherent meaning with the help of world knowledge, the interplay between the degree of agentivity of the protagonists and the case hierarchy—this interplay determines the mapping of arguments onto complements—*and* syntactic rules of valency modification (cf. 10.2.5.2; cf. the *Alternantenproblem* in Ágel 2000).

10.2.5.2 *Valency alteration*

Valency modifications such as *She forgot* (from: *She forgot to wash the car*) (reduction) and *She is washing the car with a sponge* (increase) can be addressed as *scenario conserving valency modifications* (Fischer 2003). In both cases, the basic scenarios, the forgetting and the washing scenarios, have been preserved in spite of a valency modification. However, the valency increase in (27) changes the scenario of (26):

(26)	German	Der Hund ist weggelaufen.				
		'The dog has run away.'				
(27)	German	Ihm	ist	der	Hund	weggelaufen.
		Him_{Dat}	is	the	dog	away-run

'The dog has run away while in his care.'/
'It's his responsibility that the dog has run away.'

While a running-away scenario is construed in (26), the addition of a dative complement results in a secondary "responsibility scenario" being created, the responsibility relating to the primary running-away scenario (27). The valency carrier *weglaufen* does not possess a slot for a dative complement: the responsibility scenario is imposed on the verb.

The concept of a *scenario altering valency modification* or *valency alteration* (Fischer 2003) defuses various problems of VT: the problem of readings, the problem of verb form valency, the problem of c/a-delimitation, the problem of textual valency realization. It is a theory component which enables the integration of a considerable number of constructionist phenomena into a projectionist theory (cf. the problem outline in 10.1.4).

10.2.6 Structural valency realization

Tesnière already saw clearly that the binary differentiation between valency as a potential of a lexeme and valency as textual-situational realization of this potential (Tesnière's *valence libre*) does not provide a sufficient foundation for creating a typologically adequate VT (cf. footnote 1). Rather, the realization of the valency potential is also determined by general dependency-structural properties of languages (or varieties). For example, in the Romance languages (with the exception of French) pronominal subjects are not realized in the unmarked case:

(28) Italian Cerco una casa.
Look-for-1.pers.Sg.Pres. a home
'I am looking for accommodation.'

Is the Italian subject partially optional? We argue that the question is asked on the wrong level. A realization of the pronominal subject in Italian cannot be equated with the realization of a pronominal subject in, for instance, English, as the Italian realization is emphatic:

(29) Italian Io cerco una casa.
'**I** am looking for accommodation'/
'It is me who is looking for accommodation.'

The non-realization of pronominal subjects in Romance languages is not a question of register, text, context, or situation. We are rather dealing here with a different type of valency realization, the *structural valency realization*. (28) is *not* subjectless: the verb form *cerco* functions as valency carrier *and* first complement in one! This complement realization inside the verb form has, since Pasierbsky (1981), been addressed as "micro valency" (cf. the pro-drop parameter of generative grammar).

Structurally, the "standard valency model" features two different forms of complement realization: first, as an inflectional morpheme that is incorporated into the valency carrier (micro valency) and, second, as a separate phrase (macro valency). Micro valency is not restricted to the subject (see László 1988):

(30) Hungarian Eszi.
'He/she/it eats it.'

Hungarian *Eszi* features micro-realizations of both subject and direct object, but no macro-realization.

It is tempting to assume that pronominal micro valency is the cause of the lack of pronominal macro valency in the unmarked case, but this explanation is problematic on two counts: (1) there are languages with micro-realization of the subject that also require its pronominal macro-realization (German, Islandic); (2) there are languages without micro-realizations of the first and second complement that allow pronominal non-realization of both these complements (Chinese, Korean, Japanese).

The examples show that dependency relations between micro- and macro-level can be different for individual complements and languages.[10] They can thus be used for a typological classification.

The concept of micro valency has been extended beyond subject and objects to adverbial complements (Ágel 2000: 138ff.):

(31) German Sie hat das Rad **an den Zaun** gelehnt.
'She leaned/stood/put the bike against the fence.'

(32) German Sie hat das Rad **an**gelehnt.
'She leaned/stood/put the bike (against sthg).'

(33) German Sie hat das Rad **an den Zaun an**gelehnt.
'She leaned/stood/put the bike against the fence.'

(34) German Sie hat das Rad **daran / an ihn an**gelehnt.
'She leaned/stood/put the bike against it.'

The valency carrier in (31), *lehnen*, requires an adverbial complement. For (32), it is traditionally assumed that the particle verb *anlehnen* is the valency carrier, which can optionally be combined with an adverbial complement (32 vs. 33). According to Ágel (2000: 138ff.), however, *lehnen* is also the valency carrier in (32)–(34): its

[10] Macro-realizations of complements can be seen as dependent on their micro-realizations. In Eroms's (2000) dependency grammar, the German macro-subject does not depend on the lexical verb but on the inflectional person/number-morpheme, even if this is realized on an auxiliary verb. This looks only superficially like the dependency equivalent of the inflection phrase of generative grammar: in languages without micro valency (e.g., Chinese), all complements are dependent on the lexical verb. In languages with micro-realizations of both subject and object (e.g., some Hungarian structures), the macro-subject and -object are both dependent on their respective micro-realizations, without assuming a hierarchy.

adverbial complement does not have to receive a macro-realization (32), if it is attached to the valency carrier in the form of the (separable) particle *an* (cf. *head movement*, Nichols 1986). Particle and valency carrier combine to create the verb form, i.e., the adverbial complement is incorporated into the verb. The case is, however, not completely parallel to the micro-realizations in (28) and (30) since the pronominal macro-realizations *daran* or *an ihn* in (34) are not necessarily emphatic (cf. Fischer 2003: 49f.). More strongly head-marking languages such as the North American languages feature micro-realization of a number of complements, including adverbial complements and adjuncts (cf. Nichols 1986).

10.2.7 Valency change

As was indicated in 10.1.4, VT does not just place great value on the *typological* but also on the *historical* adequacy of theory formation. Historical adequacy is stipulated in the *principle of viability* (Ágel 2001). According to this principle, any linguistic description or explanation of a phenomenon must be in agreement with the linguistic description or explanation of the history of the phenomenon in question. The theoretical apparatus should ideally be such that it can deal with (a) any stage in the development of a language, and (b) allows the integration of synchronic descriptions of historical stages of the same language into a diachronic model that has the current state of a language as its natural end point.

Since the second half of the 1970s, considerable time and effort have been invested into researching the valency arrangements of historical linguistic stages (in particular those of German, but also of other languages; cf. chapter XII in Ágel et al. (2006)). Less satisfactory is the development of diachronic modeling of valency change (summaries in Heringer 2006 and Korhonen 2006). Considerable empirical, methodological, and theoretical problems have been encountered in the attempt at applying modern concepts such as the c/a-delimitation or the identification of the valency carrier to older language stages (Habermann 2007). Habermann diagnoses, in comparison with New High German, a much stronger interaction of verb valency and verb meaning with verb-independent constructions and textual-pragmatic factors in Old and Middle High German. Also, lexical meanings have become demonstrably better defined in the history of German (Reichmann 1988: 171). These two findings allow the preliminary conclusion that older language stages have to be described using a more constructionist and less projectionist approach than is suitable for modern literal languages. A particular challenge for a *viable* theory formation consists in modeling a historically *unstable* amalgam of grammatical phenomena that require either a constructionist or a projectionist description with a *stable architecture* of constructionist and projectionist theory components.

10.2.8 Applied valency research

10.2.8.1 *Valency lexicography*

Theory formation in VT was to a considerable degree motivated by the needs of descriptive work, i.e., of lexicographic and contrastive projects. Representative examples of valency lexicography include the pioneering *Wörterbuch zur Valenz und Distribution* (Helbig and Schenkel 1969; 1991) and the *Valenz. Wörterbuch deutscher Verben* (Schumacher et al. 2004). Although German is the language that has been covered most comprehensively by valency lexicography, there are valency dictionaries for other languages, e.g., *A Valency Dictionary of English: A Corpus-based Analysis of the Complementation Patterns of English Verbs, Nouns and Adjectives* (Herbst et al. 2004).

10.2.8.2 *Contrastive valency*

A number of contrastive valency dictionaries and grammars have been compiled. We can only name a few here: *Valenzlexikon deutsch-italienisch* (Bianco 1996), *Kontrastive Grammatik deutsch-serbokroatisch* (Engel and Mrazović 1986), *Kontrastive Grammatik deutsch-rumänisch* (Engel et al. 1993), and *Deutsch-polnische kontrastive Grammatik* (Engel 1999). Three languages are constrasted in the online dictionary *The Contragram Dutch–French–English Contrastive Verb Valency Dictionary* (Colleman et al. 2004). For information on English–German contrastive verb valency, also from a typological perspective, see Hawkins (1986), Durrell (2002), Fischer (1997; 2007).

10.3 Dependency grammar

10.3.1 Contextualization

The dependency notion has a long tradition: in the West it exists implicitly in Antiquity, explicitly as *determinatio* (Boethius) since the Sixth century AD, as *dependentia* since the thirteenth century AD. The historical dependency concepts are significantly different from those of structural linguistics. We are only going to discuss the latter here. In 10.3.2, properties of a "pure" DG are introduced. 10.3.3 presents customary operationalizations of the dependency concept and also addresses problems of the notion of regent (head). The chapter concludes with a comparison of the two fundamental syntactic principles of dependency and constituency.

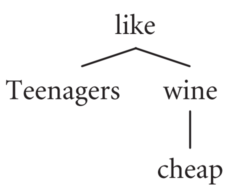
10.3.2 Dependency principles

Dependency is the second means of syntactic representation alongside constituency. A third one has not been developed (Mel'čuk 1988: 13). In practice, grammars use both, but either dependency or constituency will be favored as the more fundamental means of representation. We will now introduce basic principles of a pure DG, which define the features of a dependency representation of a sentence:

1. The elements are words (or morphemes), or more precisely word forms (or morphs).
2. Exactly one element is independent (top element of the dependency tree).
3. All other elements are dependent on an element.
4. No element is dependent on itself (no reflexivity).
5. No element is dependent on more than one other element (only downward branching).
6. If element A is dependent on element B, and element B dependent on element C, then element A is indirectly dependent on element C (transitivity).

The result of these basic principles is an actual stemma such as:

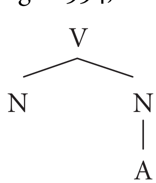
(35)



Only terminal elements occur in a dependency tree, no nodes or phrases. Each element only occurs once: the whole tree corresponds to the sentence. All elements are connected to one tree, dependency is passed down the branches. For instance, *cheap* is not only dependent on *wine* but also indirectly on *like* (but not on *teenagers*, as this word, though graphically higher than *cheap*, is not on the same branch). The basic principles impose considerable restrictions on pure dependency representations. They can only be maintained for partial descriptions of natural languages. For instance, the basic principle 5 is violated in the representation of the junction, for which Tesnière (1976: 340) himself used multiple regents.

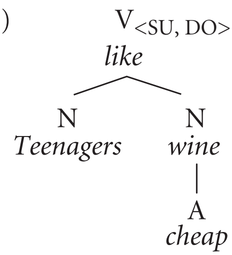
Dependency grammars are word grammars: the properties of a word that are responsible for the dependency structure are detailed in the lexicon, for instance meaning, word class, valency and other distributional restrictions. This lexical information can selectively be used to label the nuclei in a more abstract fashion (cf. Engel 1994; Heringer 1996):

(35a)



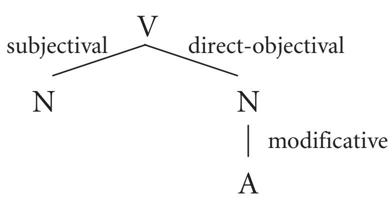
The virtual stemma (35a) also represents sentences that are structurally related to (35) such as *Children eat healthy food*. Both representations can be combined and enriched by additional lexical information:

(35b)



Finally, the dependency relations that exist between the word forms can be further specified on dependency lines (cf. Mel'čuk 1988: 16ff.; 2003: 189ff.):

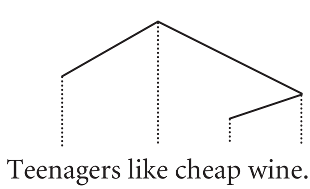
(35c)



It is not part of the system of basic dependency principles that the highest element is a verb, but in many classical dependency representations this is the case. As a result, the whole tree is a verbal node (a VP) identical with the sentence, which is thus an endocentric construction. Even if the subject (as a macro-complement) is seen as dependent on verb inflection (= the subject as a micro-complement) rather than the lexical verb (Eroms 2000), endocentricity is preserved.

The diagrams (35) to (35c) do not show the word order in the sentence. This was seen as an advantage by followers of Tesnière, since dependency representations do not aim at the linear but at the structural order of the sentence. However, dependency diagrams can be arranged in a projective manner (cf. Heringer 1996; Eroms 2000):

(35d)



10.3.3 Foundation of dependency

There are two aspects to the foundation or justification of dependency. On the one hand, there is the fundamental question of the empirical foundation and psychological reality of dependency relations. In this respect, DG has received impressive external justification (Leiss 2003). On the other hand, there is the theory internal justification of "above" and "below" in a dependency structure: how can it be determined for (structurally, i.e., non-trivially) co-occurring elements? Which of them is the regent (head) and which one the dependent? We will now introduce relation types, which were used in attempts to operationalize dependency (= D) (cf. Welke 1995):

D1 Endocentricity: the node building potential of a word category (= regent). In most cases the regent is thus a precondition for the occurrence of the dependent.
D2 Sub-categorization: the regent implies a set of dependents, determining their syntactic and semantic functions.
D3 Government: the regent determines certain formal features (case, status, canonical position, prosody).
D1 applies to optional elements in particular:

(36) Mary forgets new ← infomation quickly.

The occurrence of the noun *information* is a precondition for the occurrence of the adjective *new*. The occurrence of the verb *forget* is a precondition for the occurrence of the adverb *quickly*. In both cases the dependent can be omitted, but not the regent. D1 has problems, if both elements are obligatory and thus presume each other, e.g., *the car*[11] or *with her*. The same applies to the occurrence of *forget*: although the verb is a precondition for the occurrence of the nouns *Mary* and *information*, the opposite also applies. As English structurally requires macro-realization of subject and objects, the occurrence of the two nouns is a precondition for the occurrence of the verb *forget*. According to D1, one would have to assume bidirectional dependency (*interdependency*) between the verb and its complementation frame. But this is not permitted by the basic principles that we have introduced.

The problem is solved by D2: the verb *implies* the occurrence of the two nouns since it belongs to the sub-category of divalent verbs. This means that valency relations are used to justify the direction of dependencies: valency carriers are also governing elements in a dependency tree. But not all dependencies that occur in a sentence are implied (adjuncts, particles). Thus a dependency tree based solely on D2 would remain incomplete. D3 supports some of the dependencies established

[11] Marginally, *car* can occur by itself: *a make of car* or *Car that it is, it still doesn't drive*.

by D1 and D2, but can also contradict them ($\leftarrow_{\text{form}}$= form government; cf. 10.2.3.2 a) ii–iv):[12]

(37) French belles $\leftarrow_{\text{form}}$ filles

(38) Latin Vicino$_{\text{Dat}}$ $\leftarrow_{\text{form}}$ favet.
to-neighbor favor-(s)he.
'(S)he favors the neighbor.'

(39) Persian kûhe $\leftarrow_{\text{form}}$ boländ (Nichols 1986: 58)
mountain high
'high mountain'

(40) English The car $\rightarrow_{\text{form}}$ drives well.

Belles in (37) is according to D1 dependent on *filles* 'girls', but also according to D3, as the gender of *filles* (feminine) determines the feminine form of the adjective (*belles*, not *beaux*). *Vicino* in (38) is according to D2 dependent on *favet*, but also according to D3, as *favere* 'to favor' determines the case of *vicino* (dative). In (39) the adjective *boländ* is according to D1 dependent on *kûhe*, but the existence of an attribute has the effect that the noun *kûh* receives the suffix -*e*: thus *boländ* form-governs *kûhe*. According to D2, the noun *car* in (40) is dependent on *drives*, but the singular *car* causes *drives* to be marked as singular, thus *car* form-governs *drives*. Generally, regent or head marking as in (39) and (40) runs contrary to the dependency relations that D1 and D2 establish. If dependency relations are solely based on morphological government, the direction of the relation changes according to the morphological linguistic type (regent vs. dependent marking). Also, a purely morphological dependency notion does not establish a complete tree, since not all elements would be connected to each other (Melč'uk 1988: 109):

(41) French Il $\rightarrow_{\text{form}}$ a $\rightarrow_{\text{form}}$ dormi dans sa $\leftarrow_{\text{form}}$ chambre.
'He has slept in his/her room'

Finally, form–government relations can be complex and running in opposite directions, depending on the participating grammatical categories: in a German noun node such as *den*$_{\text{Acc}}$ *billigen*$_{\text{Acc}}$ *Wein* ('the cheap wine'), for which it is usually assumed that the determiner (*den*) is the structural regent (head) and the noun (*Wein*) is the lexical regent (core), three different form–government relations (= R) can be identified:

R1 (type: case):

grammatical structure → den billigen ← Wein

[12] We restrict discussion to (a) the (more common) notion of form government and (b) to the discussion of government that is morphologically marked. In principle, form government includes the determination of topology and prosody.

The top regent inside the phrase or node (= the internal governor) is the structural regent (head). It directly governs the lexical regent (core), indirectly the adjective. The case is externally determined: the structural regent itself is governed, its regent can be found in the respective grammatical structure. It is primarily the structural regent that is morphologically marked (*den* vs. *die, das*).

R2 (type: grammatical gender):

den → billigen Wein

The internal governor is the lexical regent (core). It directly governs the structural regent (head), indirectly the adjective. The grammatical gender is internally determined as it is an inherent grammatical feature of the noun (here: masculine). It is primarily the structural regent that is morphologically marked (*den* vs. *die, das*).

R3 (type: number):

den → billigen Wein ← intended message

The internal governor is the lexical regent (core). It directly governs the structural regent (head), indirectly the adjective. Number is externally determined, but not through government: speakers freely choose number according to the message they intend to convey. It is primarily the lexical regent that is morphologically marked (*Wein* vs. *Weine*).

According to which of the three Rs is considered, different structural descriptions result, all of which can be defended. Thus we can see that the choice of the internal governor depends on the perspective that is chosen for the description:

- If one focuses on the abstract case relation, the structural regent (head) is the internal governor.
- If, on the other hand, the abstract gender or number relation was focused on, it is the lexical regent (the noun) that is the internal governor.
- If the concrete morphological marking of the syntagmatic relations is focused on then in relation to case and gender, it is the determiner that is the internal governor, but in relation to number, it is the noun.
- Finally, if the way the respective category is determined is focused on, phrase (node)-internal determination, i.e., the gender relation, would rank highest in selecting the internal governor. Among the external factors, determination through government (case relation) would rank higher than the free choice according to the communicative intention (number relation). Accordingly, there would not just be a dichotomy (+/− internal governor (head)), but a tripartite

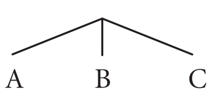
division from maximum via medium to minimum support for being the internal governor (headedness).[13]

We have discussed three possible ways of founding or justifying dependency. None of the three defines a complete dependency tree in the sense of a pure DG: D1 leads to interdependencies, D2 and D3 leave elements isolated, D3 divides into the independent relations R1–R3. In practice, dependency grammarians tend to work with multirelational dependency concepts.

10.3.4 Dependency und constituency

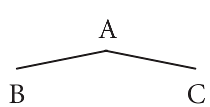
Dependency und constituency were first seen as competing means of representation. Gaifman (1965) showed that both were "weakly equivalent": chains generated by a projective dependency grammar can also be generated by a phrase structure grammar. Apart from the rather restrictive basis for the comparison, the result does not mean that dependency and constituency make identical structural claims. We will show that the opposite is the case. Our discussion draws on Engel (1994: 23ff.) and Matthews (1981: 71ff., 2007: 112ff.) (see also Fischer 1997: 13ff.; Uzonyi 2003).

Dependency and constituency are based on connexions (rule-governed co-occurrences). Constituency shows that elements that are connected by connexions form a bigger whole.

(42)


(constituency)

Dependency shows connexions directly:

(43)


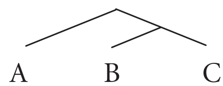
(dependency)

The *first fundamental criticism* of constituency by DG is that connexions are not shown directly but only indirectly via the constituent in which elements linked by connexions are a part. Although it is the relation between the elements that is primary and constituents are thus only derived from this relation, constituency representations turn the derived datum into the primary datum.

The constituency representation (42) is more complex than the dependency representation (43). (42) claims connexions between all three elements; a dependency representation of three elements can only show two connexions. The greater complexity of (42) is only an advantage if all three connexions in fact exist. Generally it

[13] For additional information on the structure of the German noun node, see Ágel (2006).

is a problem for constituency representations that their assumptions are too strong. Constituency representations permit further differentiation:

(44) 

(constituency)

In (44), it is claimed that A enters a connexion with the BC-grouping. In a pure DG, there is only limited scope for representing connexions between elements and groupings of elements:

(45) A (dependency)
|
B
|
C

Like (43), (45) only shows two connexions, but C is also indirectly dependent on A (principle of transitivity). Thus BC forms a node with the nucleus A. Working from the bottom of a branch upward, groupings can be read off a dependency representation. However, this is not possible if AB forms a grouping. Such a structure could be assumed for German adjective declension (cf. *den billigen Wein*), since it is both determined by the structural (head) and the lexical regent (core) (cf. R1–R3 in 10.3.3). The problem can be solved by making the additional assumption that a dependency structure can act as the nucleus of a node (cf. Lobin 1993): (A → B) → C. We can summarize our discussion up to now by stipulating that in a comparative assessment of dependency and constituency representations the achievements should be weighed against the representational cost.

The *second fundamental criticism* of constituency by DG is that constituency does not allow a natural representation of a regent (head), while DG is all about a direct representation of the regent. Compare two common constituency analyses with two common dependency representations:

(46) [[*very old*] *books*] vs. *very* ← *old* ← *books*

(47) [*no* [*fresh bread*]] vs. *no fresh* ← *bread*

It is not possible to formulate a procedure that derives the respective regents (heads) from the two constituency representations. Regents have to be identified using additional descriptive means, for instance by introducing the X-bar scheme and the dependency relations c-command and m-command and thus a convention that only allows regents (heads) to be non-phrases, while c- or m-commanded dependent elements must be phrases.

But constituency representations, too, contain information that cannot be shown in a dependency representation: one cannot derive from the dependency representation in (47) that the dependency structure of *fresh bread* itself has nuclear qualities.

We note that dependency and constituency do not merely represent connexions differently but also make different structural statements: the concept of a grouping is as alien to dependency as is the concept of a regent (head) to constituency. Therefore the two means of representation should be seen as complementary. Dependency representations might include constituency relations at certain points, while generative grammar has, with the introduction of the X-bar scheme and the relations of c-command and m-command, in effect included dependency relations in its model. These developments have somewhat defused the controversy about the better basic means of representation and point to a converging development (cf. already Schmidt 1991).

As we mentioned in 10.1.4, a different convergent development, the integration of projectionist and constructionist model components, is to be expected (see also 10.2.5 and 10.2.7). It is the syntactic theory that can solve both convergence tasks most convincingly that, in the more distant rather than the immediate future, will be able to lay claim to being the best theory.

This page intentionally left blank

CHAPTER 11

AN EMERGENTIST APPROACH TO SYNTAX

WILLIAM O'GRADY

11.1 Introduction

The preeminent explanatory challenge for linguistics involves answering one simple question—how does language work? The answer remains elusive in the face of the extraordinary complexity of the puzzles with which we are confronted. Indeed, if there has been a consensus in the last half century of work in formal linguistics, it is probably just that the properties of language should be explained by reference to principles of grammar. I believe that even this may be wrong and that emergentism may provide a more promising framework for understanding the workings of language.

In its most fundamental form, emergentism holds that the complexity of the natural world results from the interaction of simpler and more basic forces. In this spirit, emergentist work in linguistics has been pursuing the idea that the core properties of language are shaped by non-linguistic propensities, consistent with

This chapter was first written in early 2001 as a summary of my then forthcoming book (O'Grady 2005). It has been revised and updated for appearance in this volume. I am grateful to Kevin Gregg and Mark Campana for extensive discussion, to Hsiu-chuan Liao for her comments on an earlier version of this chapter, and to Miho Choo for her invaluable editorial assistance.

Bates and MacWhinney's (1988: 147) suggestion that language is a "new machine built out of old parts". O'Grady (2008*a*; 2008*d*) presents an overview of some recent emergentist contributions to the study of language.

Syntax constitutes a particularly challenging area for emergentist research, since traditional grammar-based frameworks have reported significant success in their analysis of many important phenomena. This chapter reconsiders a number of those phenomena from an emergentist perspective with a view to showing how they can be understood in terms of the interaction of lexical properties with a simple efficiency-driven processor, without reference to grammatical principles.

The ideas that I wish to put forward rest on two key claims, which can be summarized as follows:

- Syntactic theory can and should be unified with the theory of sentence processing.
- The mechanisms that are required to account for the traditional concerns of syntactic theory (e.g., the design of phrase structure, pronoun interpretation, control, agreement, contraction, scope, island constraints, and the like) are identical to the mechanisms that are independently required to account for how sentences are processed from "left to right" in real time.

The proposed unification thus favors the theory of processing, which to all intents and purposes simply subsumes syntactic theory.

A metaphor may help convey what I have in mind. Traditional syntactic theory focuses its attention on the architecture of sentence structure, which is claimed to comply with a complex grammatical blueprint. In Principles and Parameters theory, for instance, well-formed sentences have a Deep Structure that satisfies the X-bar Schema and the Theta Criterion, a Surface Structure that complies with the Case Filter and the Binding Principles, a Logical Form that satisfies the Empty Category Principle, and so on. The question of how sentences with these properties are actually built in the course of language use is left to a theory of "carpentry" that includes a different set of mechanisms and principles (parsing strategies, for instance).

I propose a different view. Put simply, there are no architects; there are only carpenters. They design as they build, limited only by the materials available to them and by the need to complete their work as quickly and efficiently as possible. Indeed, drawing on the much more detailed proposals put forward in O'Grady (2005), I suggest that efficiency is THE driving force behind the design and operation of the computational system for language.

11.2 Representations

As a first approximation, I assume that the investigation of the human language faculty requires attention to at least two quite different cognitive systems—a lexicon that draws primarily on the resources of declarative memory and a computational system whose operation is supported by working memory, sometimes called procedural memory (Ullman 2001).

I adopt a very conventional lexicon that serves as a repository of information about a language's words and morphemes, including information about their category membership (N, V, etc.[1]) and their combinatorial propensities. Thus, the entry for *drink* indicates that it is a verb and that it takes two nominal arguments. ("N" here stands for "nominal category", not just "noun".)

(1) *drink*: V, <N N>
↑ ↑
category argument grid

The computational system operates on these words and morphemes, combining them in particular ways to construct phrases and sentences, including some that are extraordinarily complex. Its operation is subject to the following simple imperative.

(2) Minimize the burden on working memory.

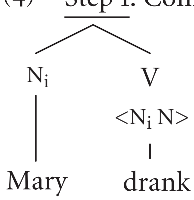
I take working memory to be a pool of operational resources that not only holds representations but also supports computations on those representations (e.g., Carpenter et al. 1994; Jackendoff 2002: 200). An obvious consequence of seeking to minimize the burden on these resources is that the computational system should operate in the most efficient manner possible, carrying out its work at the first opportunity.

(3) The Efficiency Requirement
Dependencies (lexical requirements) must be resolved at the first opportunity.

As we will see as we proceed, many core properties of English (and, presumably, other languages) follow from this simple constraint, opening the door for a memory- and processing-based emergentist account of syntax.

In forming a sentence such as *Mary drank water*, the computational system begins by combining the verb *drink* with the nominal to its left, yielding the representation depicted below. (I assume that categories are "directional"—in English, a verb looks to the left for its first argument and to the right for subsequent arguments; a preposition looks rightward for its nominal argument; and so forth.)

(4) Step 1: Combination of the verb with its first argument

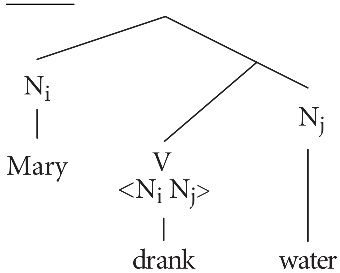


[1] I leave open the possibility that categorial contrasts are reducible to a semantic base, perhaps along the lines proposed in O'Grady (1997: 312ff.).

The resolution of an argument dependency is indicated by copying the nominal's index (representing its interpretation, as in Sag and Wasow 1999: 106–8) into the verb's argument grid. Thus, the index of *Mary* in (4) is copied into the first position of the grid of *drink* at the point where the two are combined.

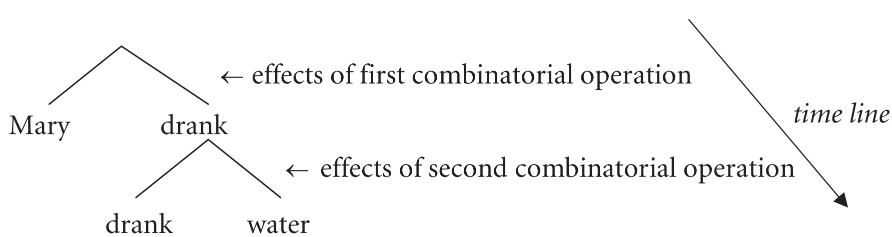
The computational system then proceeds to resolve the verb's second argument dependency by combining the verb directly with the nominal to its right, giving the result depicted below.

(5) Step 2: Combination of the verb with its second argument



Syntactic representations formed in this way manifest the familiar binary-branching design, with the subject higher than the direct object, but not because of a grammatical blueprint like the X-bar Schema. As I see it, syntactic structure is nothing but a fleeting residual record of how the computational system goes about combining words—one at a time, from left to right, in accordance with the demands of the Efficiency Requirement. Thus, the structure in (4) exists only as a reflex of the fact that the verb combined with the nominal to its left as soon as there was an opportunity to do so. And the structure in (5) exists only because the verb then went on to combine with the nominal to its right as soon as the opportunity arose. A more transparent way to represent these facts (category labels aside) might be as follows.

(6)



The time line here runs diagonally from left to right, with each "constituent" consisting of the verb–argument pair acted on by the computational system at a particular point in the sentence's formation.

11.3 Binding

Pronoun reference has long occupied an important place in theorizing about the computational system for language. The centerpiece of traditional UG-based theories is Principle A, which requires that reflexive pronouns be bound (i.e., have a c-commanding[2] antecedent), roughly in the same minimal clause. Thus, (7a) is acceptable, but not (7b) or (7c).

(7) a. The reflexive pronoun has a c-commanding antecedent in the same clause:
Harry$_i$ described himself$_i$.

b. The reflexive pronoun has a non-c-commanding antecedent in the same clause:
*[Harry's$_i$ sister] described himself$_i$.

c. The reflexive pronoun has a c-commanding antecedent, but not in the same clause:
*Harry$_i$ thinks [$_S$ Helen described himself$_i$].

In the computational system that I propose, Principle A effects follow from the Efficiency Requirement. The key assumption is simply that reflexive pronouns introduce a referential dependency—that is, they require that their reference be determined by another element. In order to see how this works, let us assume that referential dependencies are represented by "variable indices" drawn from the latter part of the Roman alphabet (i.e., *x, y, z*). Thus, the reflexive pronoun *himself* has the representation below, with the index *x* representing the referential dependency.

(8) N_x
|
himself

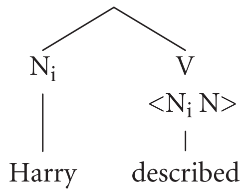
Consistent with the Efficiency Requirement, this referential dependency must be resolved at the first opportunity. But when and how do such opportunities arise? The prototypical opportunity presents itself under the following circumstances:

(9) The computational system has an opportunity to resolve a referential dependency when it encounters the index of another nominal.

Consistent with the proposal outlined in section 11.1, the computational system initiates the formation of a sentence such as *Harry described himself* by combining the nominal *Harry* with the verb and copying its index into the verb's argument grid, yielding the structure depicted below.

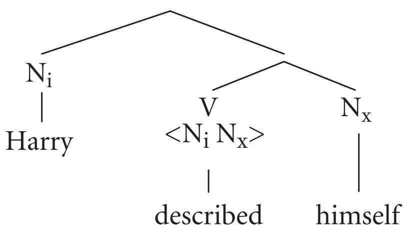
[2] X c-commands Y if the first category above X contains Y.

(10)



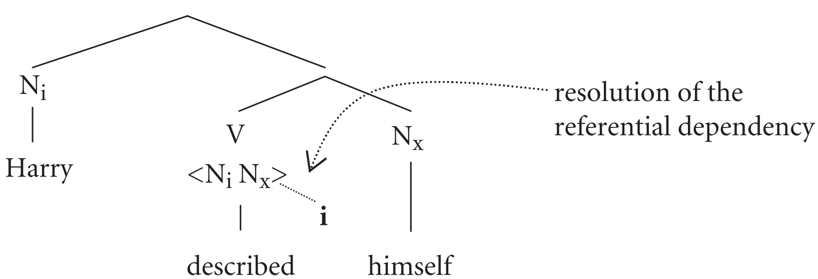
Next comes combination of the verb with its second argument, the reflexive pronoun *himself*, whose index is then copied into the verb's grid in the usual way.

(11)



This in turn creates an opportunity for the immediate resolution of the pronoun's referential dependency with the help of the index that is already in the verb's argument grid (i.e., the index of *Harry*). That is:

(12)



Given the Efficiency Requirement, no other result is possible. The verb has the opportunity to resolve its second argument dependency by combination with *himself*, so it must do so. And the reflexive pronoun has the opportunity to immediately resolve its referential dependency via the index already in the grid of the verb with which it combines, so it must do so. Anything else would be inefficient.

Now consider the unacceptability of sentences (7b) and (7c), repeated from above.

(7) b. *[Harry's$_i$ sister] described himself$_i$.
c. *Harry$_i$ thinks [$_S$ Helen described himself$_i$].

In the case of (7b), the computational system proceeds as follows.

(13) <u>Step 1</u>: Combination of *Harry* and *sister*

[Harry's$_i$ sister]$_j$

Step 2: Combination of *Harry's sister* with the verb; the index of the argument phrase is copied into the grid of the verb.

[Harry's$_i$ sister]$_j$ described
$\langle N_j\ N\rangle$

Step 3: Combination of the verb with its second argument, the reflexive pronoun *himself*; resolution of the referential dependency by the index already in the grid of the verb.

[Harry's$_i$ sister]$_j$ described himself$_x$.
$\langle N_j\ N_x\rangle$
$\Downarrow$ resolution of the referential dependency
*j

If the pronoun's referential dependency is not resolved by the index in the verb's grid in this manner, the Efficiency Requirement is violated. And if it is resolved in this way, the sentence is semantically anomalous because of the gender mismatch between *himself* and *Harry's sister*. In either case, the sentence is unacceptable.

A similar problem arises in the case of (7c). Here, the first opportunity to resolve the referential dependency associated with the reflexive pronoun arises right after the computational system combines *himself* with the verb *describe*, whose argument grid contains the index of its subject argument *Helen*.

(14) Combination of the embedded verb and its second argument, the reflexive pronoun *himself*; resolution of the referential dependency by the index already in the grid of the verb.

Harry$_i$ thinks [Helen$_j$ described himself$_x$]
$\langle N_j\ N_x\rangle$
$\Downarrow$ resolution of the referential dependency
*j

If the index of *Helen* is used to resolve the referential dependency introduced by *himself*, a gender anomaly arises. If the index of *Helen* is not used, there is a violation of the Efficiency Requirement. Either way, the sentence is unacceptable.

11.3.1 Plain pronouns

But what of plain pronouns such as *him* and *her*? In the classic Binding Theory, they are subject to a constraint (Principle B) that ensures that they cannot have a c-commanding antecedent in the same clause—hence the unacceptability of sentences such as the following.

(15) *Harry$_i$ described him$_i$.

The key observation is that there is no principled limit on the set of potential antecedents for a plain pronoun—*him* in (15) could refer to anyone who is made salient by the discourse and/or the background knowledge of the speaker and hearer. It is therefore evident that the interpretation of plain pronouns falls outside the domain of the sentence-level computational system, whose drive for quickness limits it to the consideration of "local" antecedents, as we have seen.

It is generally agreed that the interpretation of plain pronouns falls to a cognitive system—call it the "pragmatic system" for convenience—whose primary concern is discourse salience and coherence (e.g., Kehler 2002). We can represent this intuition as follows, with "→ P" indicating that the interpretation of the referential dependency introduced by the plain pronoun is passed from the sentence-level computational system to the pragmatic system for resolution.

(16) Harry$_i$ described him$_{x \rightarrow P}$

Why then can the pragmatic system normally not be used to select the salient nominal *Harry* as antecedent for *him* in (16)? Because, I propose, the pragmatic system—with its much wider range of options and its much larger domain—places a greater burden on working memory than does the sentence-level computational system, whose operation is far more locally focused. Plain pronouns are therefore COMPUTATIONALLY less efficient, and their use is shunned where the more efficient alternative—a reflexive pronoun—is available. Thus (16) is unacceptable with *him* referring to Harry simply because the same interpretation could be achieved more efficiently via the reflexive pronoun, as in *Harry described himself.* (See Reinhart 1983: 166 and Levinson 1987*b*: 410 for a similar suggestion from a pragmatic perspective.)

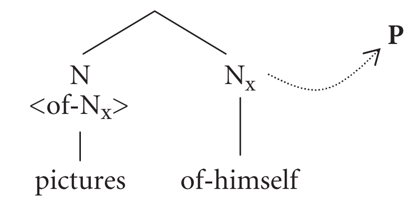
11.3.2 "Long-distance" reflexives

A long-standing puzzle for theories of pronoun interpretation stems from the fact that reflexive pronouns may sometimes take a "long-distance" antecedent. The pattern in (17) offers a typical example.

(17) John$_i$ insisted that [pictures of himself$_i$] had appeared in yesterday's newspaper.

As (18) below illustrates, immediate resolution of the referential dependency introduced by the reflexive pronoun is impossible in this case since the noun with which it combines has no other index in its argument grid. As a result, the computational system—which is compelled to act with alacrity, or not at all—passes the referential dependency to the pragmatic system for resolution there.

(18)



This creates the illusion that the anaphor is somehow "exempt" (in the sense of Pollard and Sag 1992) from grammatical principles. In fact, no grammatical principles were ever in play; the phenomenon of long-distance anaphora simply reflects the inaction of the efficiency-driven computational system.

Because the domain of the pragmatic system is far broader than that of the sentence-based computational system, the eventual antecedent of the anaphor in a pattern such as (18)—selected with attention to discourse salience—may even lie in another sentence.

(19) Antecedent outside the sentence:
Larry$_i$ had left his room in a terrible state. Pictures of himself$_i$ lay on the floor, the dishes had not been washed, the bed was unmade . . .

The fact that reflexive pronouns in contexts such as this are dealt with by the pragmatic system and can therefore be associated with a distant antecedent dramatically reduces the computational advantage that ordinarily makes them preferable to plain pronouns. This opens the door for competition between reflexive and plain pronouns, as the following example illustrates. (O'Grady 2005: 40ff. considers a much broader range of cases.)

(20) Larry$_i$ had left his room in a terrible state. Pictures of **himself/him**$_i$ lay on the floor, the dishes had not been washed, the bed was unmade . . .

Table 11.1 summarizes the contrast between the two types of pronouns in the system I propose.

In sum, there are no binding principles per se—that is, no autonomous grammatical constraints on coreference. The interpretive facts for which such principles have traditionally accounted emerge from more fundamental computational factors. As we have seen, the constraints embodied in Principle A simply follow from

Table 11.1. Plain and reflexive pronouns in English

How the referential dependency is dealt with	Type of pronoun
Immediate resolution by the computational system	Reflexive pronoun is obligatory; plain pronoun is forbidden
No opportunity for immediate resolution by the computational system; recourse to the pragmatic system	Reflexive pronoun and plain pronoun may alternate with each other

the Efficiency Requirement—reflexive pronouns are just words whose referential dependencies must be resolved at the first opportunity (immediately, if possible). And plain pronouns are just words whose referential dependencies escape the immediate interpretive action typical of the sentence-level computational system, relying instead on resolution by a pragmatic system that is sensitive to factors such as perspective and salience rather than the burden on working memory.

11.4 Control

Now let us consider the status of so-called "control structures" such as (21), in which the subject argument of the embedded verb is not overtly expressed.

(21) Harry hopes [to succeed].

The key intuition here is that there are two ways to "project" or express an argument requirement. On the one hand, it can be expressed as a categorial dependency—i.e., as a dependency that is resolved by combination with an overt nominal, as happens in the case of finite verbs (e.g., *Harry succeeded*, *Mary drank water*, etc.).

(22) V [+fin]: <N ...>

Alternatively, an argument requirement may be projected as a referential dependency (see the preceding section), as illustrated below.

(23) V [−fin]: <x ...>

This idea, which is similar in spirit to proposals found in Starosta (1988) and Sag and Pollard (1991), contrasts with the more commonly held view that subjects of infinitival verbs are expressed by PRO, a free-standing null pronoun. If we are on the right track, it should be possible to dispense with control theory, deriving its effects from more basic forces.

The two most important generalizations of traditional control theory are as follows (e.g., Chomsky 1981; Manzini 1983).

(i) The covert subject of an infinitival clause in complement position is coreferential with an argument of the immediately higher verb—with *Jean*, but not *Tim*, in the following sentence.

(24) Tim_i thinks that [Jean_j decided [$\text{PRO}_{j/*i}$ to leave]].

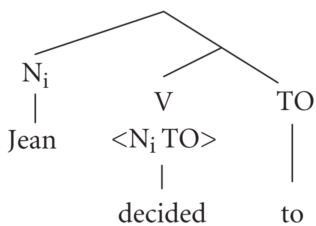
(ii) The covert subject of an infinitival clause in subject position can be interpreted pragmatically. Thus the sentence below can have the interpretations "for anyone to leave now", "for him to leave now", or "for us to leave now".

(25) Tim thinks that [[PRO$_i$ to leave now] would be impolite].

These generalizations follow automatically from the manner in which the efficiency-driven computational system seeks to resolve dependencies, namely at the first opportunity.

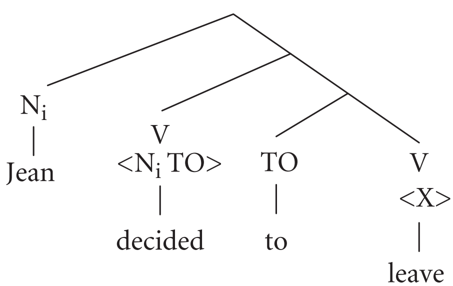
Let us begin with patterns such as *Jean decided to leave*, in which the unexpressed agent argument of the infinitival verb is obligatorily coreferential with the subject of the matrix verb. Just prior to the addition of the embedded verb, the sentence has the structure depicted below. (I assume that the infinitival marker *to* belongs to a single-member category that I will label "TO".)

(26)



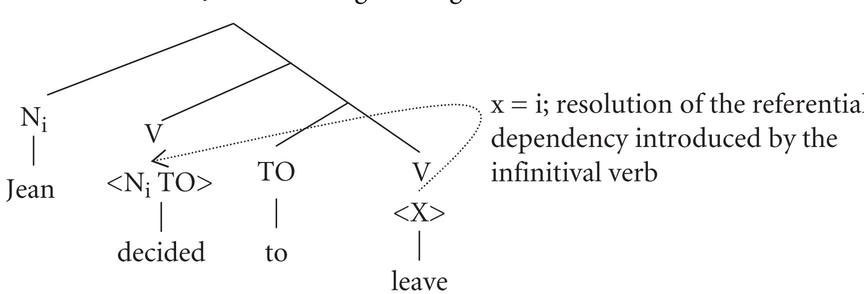
The embedded verb is then added, introducing a referential dependency (represented as x) that corresponds to its subject argument.

(27)



This referential dependency can be resolved instantly and locally, thanks to the presence of the index of *Jean* in the argument grid of the matrix verb *decide*.

(28)



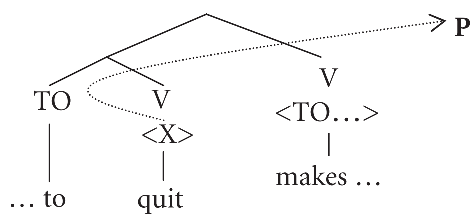
This is the only result compatible with the Efficiency Requirement; long-distance and sentence-external antecedents are thus automatically ruled out in this case.

Matters are quite different in patterns such as the following, in which the infinitival verb functions as first argument of the verb *make*.

(29) Jean said that [*to quit* makes no sense].
(="for Jean to quit now...", "for us to quit now...", "for anyone to quit now...")

As illustrated below, *make* has no index in its argument grid at the point at which it combines with the infinitival phrase. In the absence of an immediate opportunity to resolve the referential dependency associated with the infinitival verb, the dependency is transferred to the pragmatic system. This in turn opens the door for the observed range of non-local interpretations.

(30)

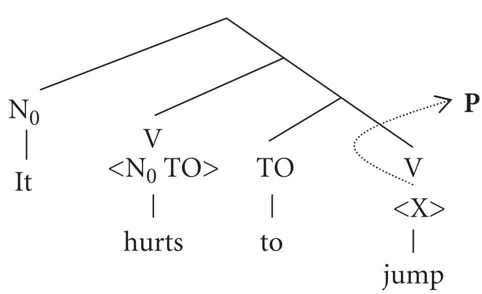


Now consider patterns such as (31), in which the infinitival combines with a verb whose only other argument is the expletive *it*.

(31) It hurts [to jump].

By definition, expletives do not have referents and thus cannot have referential indices—a property that I will represent by assigning them the "dummy" index 0. Sentence (31) therefore has the structure depicted below just after addition of the embedded verb.

(32)



Given the absence of a referential index in the argument grid of *hurt*, the referential dependency introduced by the infinitival verb cannot be satisfied by sentence-level computational mechanisms. It is therefore transferred to the pragmatic system for eventual resolution there, giving the desired generic and logophoric interpretations ("It hurts when one jumps" and "It hurts when I/you jump").

In sum, the core properties of control theory appear to follow straightforwardly from the workings of the same computational system that is used to build syntactic representations and to resolve the sorts of referential dependencies associated with reflexive pronouns. The key idea is simply that the computational system seeks to resolve the referential dependency corresponding to the subject argument of an infinitival verb at the first opportunity. As we have seen, this gives the correct result in an important set of cases: the dependency is resolved by an index in the argument grid of the matrix verb when such an index is available (as in *Jean decided to leave*) and is otherwise resolved pragmatically, resulting in the generic and logophoric interpretations observed in the examples considered above. O'Grady (2005, chapters 4 and 5) examines many other cases, including the contrast between control and raising.

11.5 Agreement

As a first approximation, English seems to require a match between a verb's person and number features and those of its subject. (For the sake of exposition, I use Roman numerals and upper case for nominal features, and Arabic numerals and lower case for verbal features.)

(33) Third person singular subject, third person singular verb form:
One remains.
IIISG 3sg

(34) Third person plural subject, third person plural verb form:
Two remain.
IIIPL 3pl

Agreement reflects the interaction of lexical and computational factors. On the lexical side, inflected verbs can introduce an "agreement dependency"—they carry person and number features that must be matched at some point with compatible features elsewhere in the sentence.

(35) a. *remains*: V, <N>
3sg

b. *studies*: V, <N N>
3sg

But how are such dependencies resolved? The lexicon is silent on this matter, and there is of course no agreement "rule" or comparable grammatical device. Rather the problem is left to the computational system to deal with—which it proceeds to do in the usual way, by resolving the dependencies at the first opportunity.

Let us assume that an opportunity to deal with agreement dependencies arises when the computational system seeks to resolve an argument dependency involving a feature-bearing nominal. In the case of a simple sentence such as *One remains* then, a chance to resolve the agreement dependency presents itself when the verb combines with its third person singular subject argument. (I use a check mark to indicate resolution of an agreement dependency. For simplicity of exposition, I do not represent argument dependencies in what follows.)

(36)

N_i	V
IIISG	3sg√
One	remains

If there is a feature mismatch at the point where the verb resolves its first argument dependency, as happens in the following sentence, the computational system faces an insurmountable dilemma.

(37) *We visits Harvey every day.
IPL 3sg

Because the presence of person and number features on the verb's first argument creates an opportunity to resolve the verb's agreement dependencies, either the computational system must bypass that opportunity, in violation of the Efficiency Requirement, or it must ignore the feature clash between the first person plural subject and the third person singular verb. Neither option can lead to an acceptable result.

The end result of all of this is that verbal agreement will be subject-oriented in all but one type of pattern. As illustrated in the following example, English verbs whose first argument is the featureless expletive *there* agree with their second argument.

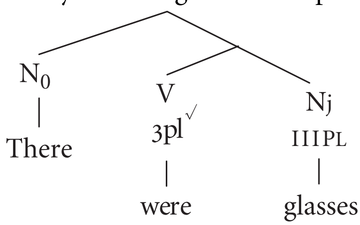
(38) a. There was **glass** on the floor.
b. There were **glasses** on the floor.

Our computational system offers a straightforward explanation for this: because the expletive *there* is featureless, it offers no opportunity for the verb to resolve its agreement dependencies. As illustrated in (39), the first opportunity to resolve these dependencies therefore arises at the point where the verb combines with its complement.

(39) Step 1: Combination with *there* resolves the verb's first argument dependency, but offers no opportunity for resolution of its agreement dependencies.

N_0	V
	3pl
There	were

Step 2: Combination with *glasses* resolves the verb's second argument dependency and its agreement dependencies.



11.5.1 Agreement and coordination

A particularly striking agreement phenomenon arises in the case of coordinate structures such as the following, where the verb can sometimes agree with the first nominal inside a conjoined phrase.[3]

(40) There **is** [paper and ink] on the desk.

The computational system builds this sentence as follows.

(41) Step 1: Combination of the verb with its expletive subject. Because *there* is featureless, there is no opportunity to resolve the verb's agreement dependencies here.

[There is]
3sg

Step 2: Combination of the verb with the first conjunct of its second argument; resolution of the agreement dependencies

There [is paper]
3sg√ IIISG

Step 3: Addition of the conjunction

There is [paper and]
3sg√ IIISG

Step 4: Addition of the second conjunct

There is [paper [and ink]]
3sg√ IIIPL

The key step is the second one, where an opportunity arises to resolve the verb's agreement dependencies with the help of the first conjunct of the coordinate noun phrase. Taking advantage of this opportunity, as demanded by the Efficiency

[3] In a survey of twelve speakers conducted by Sobin (1997), the pattern with the plural form of the verb (*There are paper and ink* ...) received a mean rating of just .81 out of 5, compared to 3.58 for the pattern with the singular form of the verb.

Requirement, results in singular agreement even though later addition of the second conjunct ends up creating a plural argument.

As expected, the singular agreement option is impossible where the first conjunct is plural, in which case the verb must carry the plural number feature in order to satisfy the demands of the Efficiency Requirement.

(42) There *is/are [papers and ink] on the desk.
3pl IIIPL

As also expected, partial agreement is possible only when the coordinate NP follows the verb. Where it appears to the left, and is therefore a fully formed plural NP before the verb is encountered, partial agreement is impossible.

(43) [Paper and ink] are/*is on the desk.
IIIPL 3pl

A variety of otherwise puzzling cases of agreement in English and other languages are considered by O'Grady (2005: 96ff; 2008*b*; 2008*c*).

In sum, the workings of verbal inflection in English reveal that efficiency, not grammatical relations, drives the agreement process. A verb agrees with its "subject" only when this NP provides the first opportunity to resolve the agreement dependencies. In cases where the subject has no person and number features, the verb agrees with its second argument—as illustrated by patterns containing the expletive *there* (*There* ***is*** *a man at the door* vs. *There* ***are*** *two men at the door*). And in cases where that NP is a coordinate phrase, we see an even more radical manifestation of the Efficiency Requirement—agreement with the first conjunct.

11.6 Constraints on *wh* dependencies

A central concern of syntactic theory involves the existence of restrictions on *wh* dependencies—the relationship between a "filler" (typically a *wh* word) and an "open" argument position associated with a verb or preposition.

(44) What did the explorers discover?

Let us assume that, like other sorts of dependencies, *wh* dependencies must be resolved at the first opportunity in accordance with the Efficiency Requirement. Furthermore, let us assume that a chance to resolve this sort of dependency arises when the computational system encounters a category with an open position in its argument grid. This is precisely what happens in the case of (44), of course, where the open argument position in the grid of *discover* creates the opportunity to resolve both the *wh* dependency introduced by *what* and the argument dependency associated with the verb.

(45) What did the explorers$_i$ discover?
<wh> <N_i N_{wh}>

It is well known that *wh* dependencies are blocked under certain conditions, including those found in "*wh* island" patterns such as (46), in which the sentence-initial *wh* word cannot be associated with the embedded clause, which begins with a *wh* phrase of its own.

(46) **What* were you wondering [*which clothes* to do with]?
(cf. I was wondering [which clothes to do something with].)

Kluender and Kutas (1993; see also Kluender 1998) suggest that the ungrammaticality of such patterns stems from the burden they create for working memory. Because holding a *wh* dependency is difficult, they argue, working memory balks at having to deal with more than one *wh* phrase per clause—as it must do in *wh* island patterns.

There must be more to it than this, however, since some *wh* island patterns are in fact quite acceptable, as (47) illustrates (e.g., Richards 1997: 40).

(47) *Which clothes* were you wondering [*what* to do with]?

This is problematic both for the Kluender–Kutas account and for standard syntactic accounts. Why should there be such a contrast?

O'Grady (2005: 118ff.) suggests that the answer may lie in how working memory stores information. One commonly mentioned possibility (e.g., Marcus 1980: 39, Kempen and Hoenkamp 1987: 245) is that working memory makes use of push-down storage—which simply means that the most recently stored element is at the top of the "memory stack" and therefore more accessible than previously stored elements.

In the case of a sentence such as (46), *what* appears first and therefore ends up being stored lower in the stack than the later-occurring *which clothes*.

(48) **What* were you wondering [*which clothes* to do with]?
a. The computational system encounters and stores *what*:

What …
MEMORY STACK
what

b. At a later point, *which clothes* is encountered and stored at the top of the stack:

What were you wondering [***which clothes*** **…**
MEMORY STACK
which clothes
what

This is the reverse of what the computational system needs for this sentence. This is because the first opportunity to resolve a *wh* dependency arises at the verb *do*, which has an open argument position corresponding to its direct object. For the sake of semantic coherence, *what* should be associated with that position (cf. *do what with which clothes*), but this is impossible since it is "trapped" at the bottom of the memory stack.

(49) ****What*** were you wondering [***which clothes*** to do with]?
memory stack
which clothes
what X

This places the computational system in an untenable position—it must either associate *which clothes* with *do*, yielding a semantically infelicitous result (cf. *do which clothes with what*), or it must spurn the opportunity to resolve a *wh* dependency, in violation of the Efficiency Requirement. Neither option is viable.

No such problem arises in the relatively acceptable sentence in (47), repeated below as (50). Here, *which clothes* is stored first and therefore ends up lower in the stack than the later-occurring *what*.

(50) *Which clothes* were you wondering [*what* to do with]?

a. The computational system encounters and stores *which clothes*:

Which clothes ...
MEMORY STACK
which clothes

b. At a later point, *what* is encountered and stored at the top of the stack:

Which clothes were you wondering [***what ...***
MEMORY STACK
what
which clothes

This is a felicitous result, since the computational system needs access to *what* first.[4]

(51) *Which clothes* were you wondering [***what*** to do with]?
memory stack
what
which clothes

The prospects for processing accounts of other island effects are excellent, and work in this area has been under way for some time (Kluender and Marta Kutas 1993; Kluender 1998; Hawkins 2004*b*; Hoffmeister et al. 2007), sometimes in combination with pragmatic analysis (e.g., Deane 1991; Kuno and Takami 1993).

[4] This gives a nested dependency in the sense of Fodor (1978).

11.7 Processing

So far, our discussion has focused on the claim that important properties of various core syntactic phenomena follow from the drive to minimize the burden on working memory, as embodied in the Efficiency Requirement. This is a necessary first step toward our goal of reducing the theory of grammar to the theory of sentence processing, but it takes us only halfway to our objective. In order to complete the task, we must establish that the computational system described here and the processor posited by psycholinguists are one and the same. More precisely, we need to show that the processor has the properties that we have been ascribing to the system that does the work of the grammar.

A defining feature of work on sentence processing is the assumption that syntactic structure is built one word at a time from left to right. As Frazier (1987: 561) puts it, "perceivers incorporate each word of an input into a constituent structure representation of the sentence, roughly as [it] is encountered" (see also Frazier 1998: 126 and Pickering and Traxler 2001: 1,401, among many others). This is just what one expects of a cognitive system that has to deal with complex material under severe time constraints. As Frazier and Clifton (1996: 21) observe, the operation of the processor reflects "universally present memory and time pressures resulting from the properties of human short-term memory". Humans, they note, "must quickly structure material to preserve it in a limited capacity memory" (see also Deacon 1997: 292–3, 331 and Frazier 1998: 125).

But what does the psycholinguistic literature say about the resolution of referential dependencies, agreement dependencies, and *wh* dependencies? Are they in fact all resolved at the first opportunity?

Nicol and Swinney (1989) make use of a cross-modal priming task to investigate the processing of English pronouns. Experiments of this type call for subjects to indicate whether they recognize probe words that are flashed on a screen at various points as they listen to sentences. The key assumption, validated in previous work, is that subjects make quicker decisions about probe words that are semantically related to words that they have recently accessed.

Now, if referential dependencies are in fact resolved at the first opportunity, as demanded by the Efficiency Requirement, the reflexive pronoun *himself* should reactivate *the doctor for the team* (its antecedent) in a sentence such as (52). This in turn should result in a shorter reaction time for a semantically related probe word such as *hospital* that is presented right after the reflexive is heard.

probe here
↓

(52) The boxer told the skier [that the doctor for the team would blame **himself** for the recent injury].

Nicol and Swinney's results bore out this prediction: probe words that were semantically related to *doctor* had accelerated reaction times after *himself*. This is just what one would expect if referential dependencies are interpreted at the first opportunity (immediately, in these patterns). More recent work (e.g., Sturt 2003; Runner et al. 2006) confirms the promptness with which the processor acts on referential dependencies.

There is also good evidence for immediate resolution of the referential dependencies corresponding to the unexpressed subject argument of infinitival verbs. A particularly promising study in this regard was carried out for Spanish by Demestre et al. (1999), who exploited the fact that the gender of the adjective *educado/educada* "polite" in the following patterns is determined by the (unexpressed) first argument of *ser* "be"—*María* in (53a) requires the feminine form of the adjective and *Pedro* in (53b) requires the masculine.

(53) a. Pedro ha aconsejado a María ser más **educada**/***educado** con
Peter has advised to Maria to.be more polite-Fem/Masc with
los trabajadores.
the employees.
'Peter has advised Maria to be more polite with the employees.'

b. María ha aconsejado a Pedro ser más **educado**/***educada** con
Maria has advised to Peter to.be more polite-Masc/Fem with
los trabajadores.
the employees.
'Maria has advised Peter to be more polite with the employees.'

Drawing on ERP data,[5] Demestre et al. found a significant wave form difference right after the adjective for the acceptable and unacceptable patterns, with the gender mismatch triggering a negative-going voltage wave. As the authors note, gender agreement errors could not have been identified so quickly if the computational system had not already interpreted the unexpressed subject argument of the infinitival verb. This is exactly what one would expect if the referential dependencies involved in control patterns are resolved at the first opportunity, as required by the Efficiency Requirement.

Agreement dependencies also seem to be resolved as promptly as possible. In an ERP study, Osterhout and Mobley (1995) had subjects read sentences such as (54) and then judge their acceptability.

(54) *The elected officials hopes to succeed.

The agreement mismatch triggered an almost immediate positive spike in electrical activity that peaked about 500 milliseconds after the violation—the usual response to syntactic anomalies on this sort of task. A similar finding is reported by Coulson et al. (1998). This suggests an attempt to resolve the verb's agreement dependencies

[5] Event-related potentials are voltage peaks that arise in the course of sentence processing and are sensitive to various linguistic factors, including agreement mismatches.

as soon as an argument carrying person and number features is encountered, as suggested in Section 11.5.

Finally, there is compelling evidence that *wh* dependencies, too, are, resolved at the first opportunity. One such piece of evidence comes from the measurement of ERPs to determine at what point speakers perceive the anomaly in the second of the following two sentences.

(55) a. The businessman knew [**which customer** the secretary called _ at home].
b. *The businessman knew [**which article** the secretary called _ at home].

If in fact the *wh* dependency is resolved at the first opportunity, then the anomaly in (55b) should be discerned right after *call*—whose open argument position should trigger action by the processor. Working with visually presented materials, Garnsey et al. (1989) uncovered a significant difference in the wave forms for the two sentences immediately after the verb, suggesting that this is indeed the point where the *wh* dependency is resolved. Evidence for early resolution of *wh* dependencies also comes from studies of eye-tracking (Traxler and Pickering 1996; Pickering and Traxler 2001) and cross-modal priming (Swinney et al. 1988). Pulvermüller (2000) discusses the neurological correlates of push-down storage, the other key processing assumption underlying my account of *wh* island effects.

In sum, there is good reason to think that the efficiency-driven computational system that is required on independent grounds to account for the defining characteristics of phenomena ranging from pronoun interpretation to agreement to filler-gap dependencies is the same system at work in sentence processing. This in turn points toward the viability of the central idea that I have been outlining: the theory of grammar should be subsumed by the theory of sentence processing.

11.8 Why aren't all languages the same?

This brings us to a crucial and difficult question. If a general efficiency-driven computational system is in fact responsible for determining how language works, then why aren't all languages the same? The answer is simple: they are—with respect to the mandate to resolve dependencies in an expeditious manner. This still leaves very significant room for variation.

For one thing, languages differ from each other in terms of the properties of their lexical items. Although the processor is compelled by internal memory-related considerations to resolve dependencies at the first opportunity, language-particular factors determine whether (for instance) the opportunity to resolve a verb's first argument dependency will occur to the left (as in English *I ran*) or to the right (as in Tagalog *Tumakbo ako* "ran I").

Or consider the subtler example of reflexive pronouns. As we have seen (Section 11.3), English reflexive pronouns simply introduce a referential dependency, which the sentence-level computational system then tries to resolve at the first opportunity. In other languages though, so-called "reflexive pronouns" carry more than just a referential dependency; they are used to encode subtle pragmatic contrasts involving logophoricity and perspective, as Huang's (1994) detailed study of the Mandarin system demonstrates. These are not the sorts of factors that the computational system is able to deal with. The interpretation of such pronouns must therefore be passed to the pragmatic system, whose priorities lead to a very different type of syntax, as illustrated in the following examples from Kao (1993: 37–8).

(56) a. "Reflexive pronoun" with a non-local antecedent:
Sam$_i$ renwei [Lisi$_j$ bu yinggai piping ziji$_{i,j}$].
Sam think Lisi not should criticize self
'Sam$_i$ thought that Lisi$_j$ should not criticize self$_{i,j}$.'

b. "Reflexive pronoun" with a non-c-commanding antecedent:
Lisi-de$_i$ jiaoao hai le ziji$_i$.
Lisi-Gen pride hurt Asp self
'Lisi's$_i$ pride hurt self$_i$.'

Languages may also vary with regard to the size of the burden on working memory that they are willing to tolerate. The syntax of *wh* dependencies offers a revealing example of how this works. As Hawkins (2004*b*: 192ff.) documents in considerable detail, languages differ in a principled way with respect to acceptability of filler-gap relationships that extend across a clause boundary. Some languages (e.g., Russian) permit extraction only from reduced (infinitival) embedded clauses.

(57) the cucumber which I promised [$_{INF}$ to bring _]

Others (such as English) allow extraction from more elaborate tensed embedded clauses.

(58) the cucumbers which I promised [$_S$ that I would bring _].

And still others (e.g., Swedish) tolerate extraction from a tensed clause that is embedded inside an NP.

(59) a bone which I see [$_{NP}$ a dog [$_S$ which is gnawing on _]]

The crucial generalization seems to be this: if a language permits the computationally more demanding pattern, then it must also permit its less difficult counterparts. Thus, English permits not only (58) but also (57), and Swedish allows not just (59) but also (57) and (58).

This state of affairs is reminiscent of what we find in phonology, where it is recognized that considerations relating to ease of production and perception motivate a variety of processes—word-final devoicing, vowel harmony, intervocalic voicing, and so forth. At the same time, there is no expectation that all languages will have intervocalic voicing or manifest vowel harmony. Rather, it is understood that the effects of articulatory and perceptual pressures are manifested via an implicational logic—if the more "difficult" sound is permitted in a particular position, then so is its less difficult counterpart. Thus if a language allows voiced obstruents in word-final position (English, but not German), it must also permit voiceless obstruents in that position. Comparable asymmetries, motivated by processing difficulty, are pervasive in syntax and make up a good deal of the principled variation found across languages.

11.9 The problem of language acquisition

Space does not permit an adequate discussion of language acquisition, which is seen by many as the premier explanatory challenge for linguistics, but a few brief remarks are in order.

For many years, the field has been split over the question of whether children's early experience with language provides enough information to support induction of the knowledge needed to speak and understand a language (the so-called "poverty of stimulus" claim). Essentially without exception, opponents of Universal Grammar hold that the input suffices, while proponents of UG adhere to the opposite position, maintaining that without access to innate grammatical principles it would be impossible to learn the intricate details of language.

I disagree with both views. As explained in detail in O'Grady (2005; 2008*a*), I do not believe that induction from experience can be the whole story: the facts are simply too complex, the input too sparse, mastery too rapid, and errors too infrequent. On the other hand, I do not see UG as the answer either. Rather, I hold that the gap between experience and the intricacies of language is bridged with the help of the processor, which directs learners to particular options that are not evident from information available in the input—the association of reflexive pronouns with antecedents in particular positions, intricate patterns of agreement, certain scopal preferences (O'Grady 2008*a*), and so on.

The key idea is simply that children are born with an efficient brain—that is, a brain that seeks to carry out computations in a way that minimizes the burden on working memory. Early in life, language learners are exposed to adult utterances that contain lexical items with particular properties arranged in a particular order. As input of this sort is encountered and processed over and over again,

Table 11.2. Some efficiency-driven computational routines

Routine	Effect
Resolve the referential dependency introduced by *X-self* at the first opportunity	*himself=John*, not *Bob*, in *Bob thinks John mistrusts himself*
Resolve the referential dependency introduced by an infinitival verb at the first opportunity	A local controller in patterns such as *Mary decided to leave*
Resolve a verb's agreement dependencies at the first opportunity	Agreement with the first argument in *A man is here*, and with the second argument in *There is a man at the door*

routines consisting of particular computational operations and series of operations develop. Language acquisition takes place as these routines are automatized through repeated use. (As noted by Paradis 2004: 28 and Bybee and McClelland 2005: 382, a side effect of this automatization is that processing is speeded up, helping to give language use its characteristic effortlessness.)

Some computational routines, such as the ones summarized in Table 11.2, are direct instantiations of the propensity to resolve dependencies at the first opportunity. As such, they do not need to be learned per se—they are simply associated with the appropriate lexical item(s). We therefore expect more or less error-free development, which is in fact the case for the phenomena in question (O'Grady 2005: 193ff.).

Other routines develop in response to language-particular facts that are neutral with respect to computational burden. For example, given repeated exposure to the word order of English, children will develop routines in which determiners look to the right for their nominal argument, while verbs look to the left for their first argument and to the right for their second argument, and so on. Table 11.3 lists some simple routines of this sort.

As computational routines of both types become increasingly fixed and rigid over time, they restrict the options available to the processor when it operates on the words and morphemes of English. This gives a result identical in its effects to having a grammar that imposes a locality requirement on reflexive pronouns, has a subject

Table 11.3. Some language-particular computational routines

Routine	Effect
Looking to the right for the noun associated with a determiner	*the book*, not **book the*
Looking to the left for a verb's first argument	*John ran*, not **Ran John*
Looking to the right for a verb's second argument	*Eat the apple*, not **the apple eat*

agreement rule, orders determiners before nouns, places verbs before direct objects, and so forth. But what appears to be a grammar is just a processor that has become set in its ways.

If this processor-based view of language acquisition is on the right track, then we would expect to find independent evidence for processing effects in the way in which language development unfolds. This matter is discussed at some length in O'Grady (2005; 2008*a*), and I will draw briefly on that discussion here to consider the case of pronoun interpretation.

As explained in Section 11.3, the interpretation of reflexive pronouns in English follows directly from the Efficiency Requirement—the choice of antecedent is determined in most cases simply by the need to resolve the referential dependency at the first opportunity. Given that efficiency is an inborn computational imperative, we predict early mastery of reflexive pronouns—allowing of course for the time needed to identify the lexical items that function in this capacity. Plain pronouns have the potential to be somewhat more troublesome, however, since their interpretation falls outside the domain of the sentence-level computational system and requires a sensitivity to non-local factors such as discourse salience, coherence, and so forth. Indeed, it is known that the interpretation of plain pronouns is associated with increased processing difficulty in at least some cases (Sekerina et al. 2004).

The developmental facts are also very intriguing in this regard. As established in a variety of comprehension studies (e.g., Wexler and Chien 1985), children typically interpret reflexive pronouns correctly more than 95% of the time from a very early age (as young as age 3). In contrast, performance on plain pronouns during the same period is usually significantly lower, hovering between 50% and 85%, depending on how carefully the supporting contexts are constructed (Conroy et al. 2008). Moreover, typical errors involve interpreting the pronoun in patterns such as *Donald Duck washed him* as coreferential with the subject—i.e., as if it were a reflexive pronoun. This suggests an initial preference for referential dependencies that can be resolved at the least cost—through the mediation of sentence-level mechanisms that are sensitive to immediacy and locality. This is just what we would expect if the language faculty is shaped by the need to minimize the burden on working memory.

11.10 Does grammar exist?

The theory that I have been describing is emergentist in the sense outlined at the outset: it seeks to attribute the defining properties of language to more basic non-linguistic forces, particularly efficiency-related processing considerations. Such an approach raises the question of whether there is still a need for the type of cognitive system that we traditionally think of as a grammar. Two clarifications are in order.

First, I have no argument with the existence of a lexicon that contains information about category membership and selectional properties—in fact, I make use of just such a lexicon. My doubts about grammar pertain to the character of the computational system that is responsible for the many aspects of a sentence's form and interpretation that do not follow from lexical properties—the fact that reflexive pronouns require antecedents in particular positions, that verbs exhibit particular patterns of agreement, or that certain types of *wh* island patterns are unacceptable, for example.

The traditional view is that the rules and principles that regulate these phenomena must be distinguished from processing mechanisms. As Jackendoff (2002) observes, grammatical rules describe patterns of elements (p. 57); they say what the structure is, not how it is built (p. 31). According to this view, the grammar and the processor interact, but they are not the same thing (Fodor 1989: 177ff.; Frazier and Clifton 1996: 9, 25; Frazier 1998: 126). Moreover, it is widely held that the processor is subordinate to the grammar: "the most basic assumption about the nature of the human sentence processor", Ferreira et al. (2001: 13) write, is "that it obeys the fundamental principles of grammar when constructing interpretations". All of this is of course incompatible with the view I have adopted.

A second clarification has to do with the two very different senses in which the term "grammar" is used by those committed to its existence. Sometimes, "grammar" is used to refer to the putative system of inborn constraints that gives the human language faculty its unique character (i.e., Universal Grammar). And sometimes it is used for the system of rules that account for the particular patterns and facts that distinguish one language from another (i.e., "language-particular grammars", as when someone talks about writing the grammar of Balinese or the grammar of Turkish).

My position with respect to UG is that it does not exist; a simple efficiency-driven linear processor lies at the heart of the human language faculty and carries the burden of explaining why language has the particular properties that it does and how those properties are acquired with such apparent ease by children. As outlined in earlier sections of this chapter, core properties of sentence structure, binding, control, agreement, extraction, and other phenomena that have long been offered as prima facie evidence in support of UG seem to follow from the processor's commitment to minimizing the burden on working memory by resolving dependencies at the first opportunity.

My position with respect to language-particular grammars is that they, too, do not exist, but here the point is a subtler one, since many "low-level" processing routines embody the sorts of generalizations that could just as easily be stated by a grammatical rule: determiners look to the right for a noun, verbs look to the left for their first argument, and so forth. But processing routines are not just grammatical rules under another name. They are real-time PROCESSES, not static statements about how elements are arranged. Moreover, they are independently required—speech, after all, has to be processed. Rules, if they exist, must do something more than restate facts about language that follow from the operation of the processor. Fodor puts it this way:

> ... there *must* be psychological mechanisms for speaking and understanding, and simplicity considerations thus put the burden of proof on anyone who would claim that there is more than this. To defend the more traditional view, what is needed is some sign of life from the postulated mental grammar. (Fodor 1978: 470)

I see no such signs.

The proposals that I have outlined in this chapter barely merit consideration—coverage is offered for only a tiny body of data, attention has been focused on a single language, apparent counterexamples abound, and so forth. But the central thesis, however ill fated it may be, is perhaps at least clear: a single efficiency-driven computational system offers a solution to the classic problems confronting both the theory of grammar and the theory of sentence processing. A sentence's design reflects the way it is built, not the other way around—there are no architects, just carpenters.

This page intentionally left blank

CHAPTER 12

FORMAL GENERATIVE TYPOLOGY

MARK C. BAKER

12.1 Introduction

The approach to linguistic research that I describe here can be called "Formal Generative Typology" (FGT). It does not have an extensive body of official doctrine, its own entrenched name, distinctive modes of representation, or a special choice of data. It is a relatively opportunistic approach—ready to borrow insights from other linguistic approaches in the course of pursuing its goals.

Its distinctive content comes primarily from its goals and methods. FGT is firmly rooted in the Chomskyan tradition of linguistics, as that has developed over the past 50 years (Chomsky 1957; 1965). It grows most directly out of the "Principles and Parameters" stage of that tradition, which crystallized in the 1980s when Chomsky and his followers started to turn their attention more seriously to issues of crosslinguistic comparison (Chomsky 1981; Chomsky and Lasnik 1993). It is in this sense that the approach is "formal" and "generative". Yet as this line of research has been pursued in its own terms, it has come to a degree of convergence of interest and method with linguistic typology as practiced by linguists like Joseph Greenberg (1963), Bernard Comrie (1981), and others. That is why it is "typology".

To some contemporary linguists, the term "formal generative typology" might seem like an oxymoron. It is often thought that typology can only be properly pursued from a functionalist perspective, and that formal-generative techniques are inherently narrow and hard to generalize to new languages. In this chapter, I seek to show that this need not be so. In so doing, I lay out the basic goals and concepts that make FGT a fruitful and consistent enterprise, and present a few examples that illustrate its hopes and its promise.

I confess that this is a somewhat personal vision. Although virtually all of its important ingredients are borrowed from other linguists, it may be that I am the only one who tries to do exactly what is described here. Nevertheless, I believe that many other talented linguists have research programs that overlap significantly with what I describe, and that they have compatible if not identical visions. And anyone is welcome to join the enterprise to whatever degree they feel inspired to do so.

12.2 Some key questions for linguistic research

The central questions that FGT seeks answers to are the following:

(1) What properties of natural human languages are genuinely universal, inherent to the human species as such?

(2) What properties of natural human languages vary from one human language to another?

(3) Which aspects of variation are patterned, systematic, and grammatical in nature, and which aspects of variation are random, idiosyncratic, and lexical in nature?

To these three, we can add a fourth question, usually not a focus of direct inquiry, but always on the horizon, giving the enterprise its possible grander significance:

(4) What do the answers to (1)–(3) imply about the nature and origins of the human mind, of which language is a part and a reflection?

In this particular chapter, I discuss these questions only as they apply to the *syntax* of human languages, where syntax is construed broadly as including the entire compositional system by which complex expressions are built up out of simple morphemes and words. Thus defined, syntax includes not only syntax proper but also productive morphology and those aspects of semantics that involve composing and calculating linguistic representations internal to the mind of the language user.

This limitation is purely a practical one. There is no reason not to think that these four questions are equally applicable to the domains of phonology and the lexicon. However, pursuing them in those domains is outside my expertise.

There is nothing especially novel about these four questions. Any linguist could ask them, and many do. They can, for example, be seen as elaborations of questions that Chomsky has often articulated. For example, Chomsky and Lasnik (1993) state the following as the first two questions on their list of "classical problems in the study of language":

(5) a. What does Jones know when he has a particular language?
b. How did Jones acquire this knowledge?

Questions (1)–(3) can be derived from these "classical" questions simply by varying Jones and his linguistic background—in other words by comparing the knowledge Jones has of his language and how he acquired it with the knowledge that Li has of her language, and how she got it, and with the knowledge that Mohammed has of his language, and how he got it. Those bits that are shared for all choices of Jones/Li/Mohammed will constitute the answers to question (1). Those bits that differentiate their languages from one another will be the answers to question (2). And for each element in the answer to question (2), the further issue arises as to whether it is an elemental difference or a difference that derives from some other, more fundamental difference. In other words, the question arises as to whether people's knowledges of language differ in many little ways or in a few big ways. This is question (3).

How does FGT relate to the Minimalist Program? It is easy to get the impression that Chomskyan linguistics largely abandoned the quest defined by (1)–(3) in the switch from Principles and Parameters to Minimalism, initiated by Chomsky (1993; 1995) (see also Boeckx, this volume). But I think that there is no deep or meaningful conflict between them. Chomskyan linguistics (like other approaches) has always sought to maximize explanation. A simple way to think about degrees of explanation is that explanation is maximized when the ratio of observable phenomena analyzed to theoretical assumptions made is as high as possible. How then might one strive to increase the degree of explanation achieved by (say) the Principles and Parameters theory of *circa* 1990? There are two obvious strategies to pursue. The first is to reduce the number of assumptions made by the theory, while keeping the phenomena analyzed (ideally, more or less) constant. This is the spirit of the Minimalist Program proper. The second strategy is to increase the number of observable phenomena being analyzed, while keeping the number of assumptions (ideally, more or less) constant. This is the driving force of FGT. Chomsky himself has concentrated primarily on the first approach over the last 15 years, as have quite a few others; hence the Minimalist Program. In contrast, I (for one) have concentrated on the second approach. In practice, the two kinds of projects look rather different when it comes to their day-to-day activities.

But there is no principled conflict between the two research thrusts; on the contrary, they should be complementary, even synergistic. And if one looks not so much at Chomsky's own output but at that of Ph.D. students coming out of MIT and allied departments, much of it is in practice making a contribution of the second kind: people are seeking to apply a body of shared theoretical assumptions to Tagalog, or Itelmen, or Ewe, or Passamaquoddy, or whatever.

12.3 How the questions can be pursued: Borrowing the best of both

In recent linguistic history, there have been two somewhat different senses of the term "universal", and hence two rather different approaches to finding universals. It is easy for the two to misunderstand each other. A hope of FGT is that these two approaches can be found to be complementary, rather than in conflict with each other or irrelevant to each other.

12.3.1 Universality and the needs of language acquisition

Chomsky's own answer to question (5b) has famously been built on the assumption that there is a fairly rich initial state of the language faculty, which is what he calls "Universal Grammar" (UG). Moreover, this "Universal Grammar" is thought, in point of fact, to constitute a rather high percentage of the answer to (5a) as well; this is Chomsky's famous nativism. The word "Universal" in "Universal Grammar" is thus used in a somewhat counterintuitive sense. When one first hears this term, it is natural to think that it refers to rules or principles of grammar that are part of all human languages. But that is not Chomsky's primary meaning. More properly, Chomsky's Universal Grammar is the preprogrammed biases that a human being brings to bear on the task of learning any natural language. Once this terminological point is realized, Chomsky's scandalous, counterintuitive, ethnocentric sounding claim that "A great deal can be learned about UG from the study of a single language" (Chomsky 1981: 6) loses most of its scandalous character. Rather, it is virtually a truism.

To make the discussion more concrete, consider the following simplified-but-not-entirely-misleading example. It is well known that English is a rather typical example of a Subject–Verb–Object language, a word order shared by some 35% of languages of the world (Dryer 2005*b*). As such, not only does the verb come before the object but the auxiliary verb also comes before the main verb, and the directional adposition comes before its object:

(6) John will (quickly) give a book to Mary.

Moreover, standard generative (and structuralist) techniques show that the relationship of the verb to the object in English is also different from the relationship of the verb to the subject in a subtler way, that goes beyond the obvious difference in linear order. The verb and the object join together to form a constituent—the verb phrase—which does not include the subject. One reflection of this is the fact that the verb and the object are adjacent in sentences like (6), whereas the verb and the subject need not be; rather the subject can be separated from the verb by the tense auxiliary or by an adverb. Other evidence confirms this. For example, the verb and the object can be omitted from a sentence leaving the subject and tense marker behind, but the verb and the subject cannot be omitted, leaving the object and the tense marker behind:

(7) a. John will read the book tomorrow, and Mary will – too.
b. *John will read the book tomorrow and – will – the newspaper too.

Similarly, the verb and the object can shift together to the beginning of a sentence in special contexts, whereas the verb and its subject cannot do this:

(8) a. Mary promised that she would read the book, and [read the book] Mary will.
b. *Mary promised that she would read the book, and [Mary read] will the book.

For current purposes, we state this familiar fact about English in the following way:[1]

(9) *The Verb–Object Constraint (VOC)*: The verb combines with its object to form a linguistic unit that cannot include the subject.

The Chomskyan might then go on to observe that some of the empirical evidence in favor of the VOC is rather subtle. Students of a beginning linguistics class often have no firm intuitions about which is more closely connected with the verb in English, the subject or the object. In readily observable linear terms, the subject is often as close to the verb as the object is. Moreover, it is not hard to imagine that some of the children who learn this property of English do so without ever hearing sentences like those in (8), although these were important to establishing the existence of the verb phrase in the minds of linguists. How then do they acquire the knowledge expressed in the Verb–Object Constraint? The Chomskyan

[1] For current purposes, we can say that the object of the verb is the nominal that denotes the entity whose state or position changes the most in the event described by the verb, whereas the subject of the verb is the nominal that denotes the entity that caused the event to take place—although this is a simplification.

might conjecture that this is an expectation that they bring to the task of language learning, that they are predisposed to this hypothesis, and do not need a lot of conclusive evidence to arrive at it. In other words, (9) might be part of Universal Grammar.

Given this perspective, and the practices that have grown out of it, it is easy to think that the Chomskyan approach to linguistics has very little to do with linguistic typology as it is usually practiced. Nevertheless, one can draw a link between the two, given a few plausible assumptions. First, suppose that Chomsky and others are correct that the Universal Grammar that forms the basis of the answer to question (5b) also constitutes an important part of the answer to question (5a). Second, suppose that language learning is a more or less monotonic process. In other words, Jones the language learner *adds* knowledge to her innate store to arrive at full knowledge of English, but she does not discard or radically change what she already has. Third, suppose that the initial state of the language faculty is essentially the same for all human children.[2] Then, by parity of reasoning, we expect that the same UG that is a substantial part of Jones's knowledge of English is also a substantial part of Hiro's knowledge of Japanese, and of Mohammed's knowledge of Arabic, and of Sak's knowledge of Mohawk. And from this it follows that all human languages must be similar in these respects.

We can see how one is led to this by giving the example above a crosslinguistic dimension. The Nigerian language Edo, although historically unrelated to English, has the same word order as English in all relevant respects, as shown in (10).[3]

(10) Ozo ghá rhié èbé ne Emeri.
Ozo FUT give book to Mary
'Ozo will give a book to Mary.'

Chomskyan reflection on the logic of language acquisition leads us to think that children must arrive at the VOC in (9) from their innate predispositions, plus at most observations of data like (6), since there is no guarantee that they will observe

[2] This assumption is not logically necessary, but we have good observational reasons to believe it. It is logically possible that the innate assumptions about language that different children start with are significantly different (they have different "UGs"), perhaps as a result of genetic variation. But we observe that—barring serious genetic defects—any child can learn any natural human language perfectly if it is raised in the right environment, regardless of (for example) its genetic background. For example, a child with Chinese genes who is brought up in an integrated suburban American environment acquires an English that is indistinguishable (up to normal idiolectal variation) from the English acquired by a child with Western European genes growing up in the same neighborhood. Furthermore, as far as we know, they learn the language in essentially the same way, passing through the same kinds of intermediate stages.

[3] The abbreviations used in this chapter are the following: 2sO, second person singular object agreement; 2sS, second person singular subject agreement; 3sO, third person singular object agreement; 3sS, third person singular subject agreement; ACC, accusative case; ADV, adverb; DUP, duplicative; FACT, factual mood; FGT, Formal Generative Typology; FUT, future tense; HAB, habitual aspect; IMPF, imperfective aspect; IND, indicative mood; M, masculine; N, neuter; NOM, nominative case; PAST, past tense; PRT, particle; PUNC, punctual aspect; STAT, stative aspect; UG, Universal Grammar; VOC, the Verb–Object Constraint; WALS, the *World Atlas of Language Structures*.

examples like (7) and (8). But then the child exposed to Edo is in the same position as the child exposed to English—even though there is no process of VP fronting or VP ellipsis in Edo. Therefore, we expect that the object and the verb form a constituent that excludes the subject in Edo as well. And further investigation shows that this is correct—even though the evidence that is available to show that this is correct in Edo is rather different from the evidence available in English (Stewart 2001).

12.3.2 Universality and observing diverse languages

Now, what about languages in which the basic word order of subject, object, and verb are not the same as in English? The form of argument just sketched is more general, and does not necessarily depend on there being similarities of word order. Its upshot is that if there is a Universal Grammar in Chomsky's language-acquisition-oriented sense, one expects that there will be observable universals of language in something like Joseph Greenberg's sense as well.

Greenberg (1963) famously initiated the search for facts about grammatical patterning that are observably true in representative samples drawn from the set of all natural languages. These are patterns that recur in languages that are not (recently) historically related—universals of language in a more obvious sense. (11) is a classic example of this sort of universal:

(11) **Universal 4**: With overwhelmingly greater than chance frequency, languages with normal Subject–Object–Verb order are postpositional.

Consider, for example, Japanese, a canonical positive example of Greenberg's Universal 4. The Japanese equivalent of English (6) or Edo (10) is (12).

(12)	John-ga	Mary-ni	hon-o	yat-ta.
	John-NOM	Mary-to	book-ACC	give-PAST

At first glance, the differences between Japanese and English/Edo are more striking than the similarities. In particular, the position of the verb relative to the object is different, the position of the adposition relative to its associated NP is different, and the position of the auxiliary verb with respect to the main verb is different. This is the other very common word order, found in some 40% of languages of the world (Dryer 2005*b*).

Now, given this salient difference in word order, what do we make of the conjecture that the VOC is part of Universal Grammar, an innate bias for language acquisition? If it is, then the VOC will influence the way that a Japanese child learns Japanese, too, and hence it will influence the structure of mature Japanese. And indeed there is reason to believe that this is true. First, despite the different word order, Japanese is like English and Edo in that the direct object is adjacent to the verb in the simplest, least marked word order, whereas the subject is not.[4] Moreover, this phrasal grouping can be confirmed by other, less obvious tests. For example, Japanese has a process similar to VP anaphora in English, in which the verb and the object may be omitted, but the subject remains (example from Tsujimura 2006):

(13)	Taroo-ga	tori-o	uti-korosita-node	Hanako-mo	soo-sita.
	Taro-NOM	bird-ACC	shoot-to-death-because	Hanako-also	do.so
	'Since Taro shot a bird to death, Hanako did so too.'				

So the generative conjecture that the VOC is universal holds up rather well. Indeed, the Greenbergian universals tacitly presuppose that head-final languages like Japanese and head-initial languages like English have essentially the same phrasal groupings across the board (see Baker 2001: ch. 3 for discussion).

Looking more broadly, it is consistent with the VOC that the two most common word orders in the world, together accounting for more than 75% of all languages, both have the object adjacent to the verb. The VOC thus captures something important that those two word orders share that might help account for their popularity. The VOC also has the virtue of playing a role in accounting for both obvious facts of word order and less obvious properties (like possible ellipses) in the same terms.

More generally, then, a core interest of FGT is evaluating whether this sort of convergence between the two sorts of universal is found in general or not. We want to see if it is true in practice that what one is led to attribute to Universal Grammar from the detailed study of (say) English, driven by the need to account for how its details could have been learned from simple and unsystematic data, is really the same as what one is led to attribute to Universal Grammar from the detailed study of Japanese, or Mohawk, or any other language. In this, the conception of what language is that informs FGT is solidly Chomskyan, but the data considered and the method used has something important in common with Greenbergian typology, with its emphasis on collecting data from a wide sample of the world's languages and looking for patterns that recur in unrelated languages. Only when we find such patterns can we say with confidence that grammar is truly universal.

12.3.3 Universals and abstractness

But despite the important point of similarity, there is still a methodological difference between FGT and the functionalist-oriented practice of typology. A crucial issue is the level of theoretical abstraction at which the search for universals and patterned variation takes place. Perhaps the most constant feature of functionalist typology in the tradition of Greenberg, Comrie, and others over the last 50 years

[4] Of course, both Japanese and English also allow marked word orders for special pragmatic purposes, and we must abstract away from this. There has been much discussion of how to identify basic word orders in both the functionalist-typological and generative literatures.

is that it focuses on superficial features of languages—features that can easily be described in the relatively atheoretical terms of traditional descriptive linguistic practice. For example, Bickel (2007: 242) writes, in describing the current typological scene:

But not everything has changed: most prominently, as in the past, typologists find it useful to develop variables as close to observable data [operationalized criteria] as possible and close to fieldwork. This is first of all a practical decision, because very abstractly defined variables are difficult to survey in sufficiently large samples, and samples can often only be completed by doing additional fieldwork. But the decision is also theoretically motivated because the definition of abstract variables is also tied to some UG model that itself seeks to abstract away from linguistic diversity, and less so to the kinds of anthropological hypotheses of interest.

This commitment to studying nonabstract, surface-observable properties of language is one of the most constant features of typology as it is usually practiced, a near corollary of its defining interest in studying representative samples of unrelated languages. As Bickel acknowledges, it is a commitment with practical motivations. The notions it uses are usually fairly easy to define, they can often be found in standard descriptive grammars, and the results can be replicated by others (in principle, anyway). But Bickel also correctly points out that this is a theoretical choice as well as a practical one. It amounts to a denial, implicit or explicit, of the value and reality of more abstract generative concepts—concepts such as phrase structure, c-command, "movement", and the like. Functionalist typology is often motivated by a kind of positivistic empiricism, which wants to see the theory emerge from the data, rather than having a theory imposed onto the data (see, for example, Croft 2002). It is an attempt to avoid the arcaneness, the question-begging, and the immunization from counterexamples that generative abstractness at its worst makes possible. And I can easily understand why one would want to avoid these things.

But it should also be acknowledged that traditional typology has in a sense failed, and thus has been led to change its goals.[5] It has by and large been unable to discover many interesting universal properties of natural human languages. At most, standard typologies find statistical tendencies of various degrees of strength. Thus, Bickel (2007: 245) also writes:.

Large datasets almost invariably reveal exceptions to universals, and this, together with a substantial increase of newly described languages and assisted by prominent conceptual argumentation (e.g., Dryer 1998, Croft 2002, ch. 8), has practically done away with notions of absolute universals and impossibilities. Modern studies of typological distributions involve statistical methods, from association tests to multivariate scaling methods .... The general

[5] Of course, many researchers within the paradigm (such as Bickel) would describe this not as a failure but as a healthy discovery that moves the field forward toward greater truth and enlightenment.

assumption is that if there are large-scale connections between linguistic structures, or between linguistic structures and geography, they consist in probabilistic (and therefore exception-ridden) correlations between independently measured variables.

This retreat from a universalist vision is also borne out in the massive, wonderful, and frustrating *World Atlas of Language Structures* (*WALS*) (Haspelmath et al. 2005), which maps out some 140 linguistic features over hundreds of languages each. Haspelmath (personal communication) has observed that computational techniques have been used to systematically mine through the work, looking in a purely mechanical way for statistically significant correlations among the properties cataloged (Cysouw et al. 2007). The result was that very few correlations emerged, and most of the ones that did looked crazy, correlating (say) some phonological feature with some syntactic feature, which no plausible linguistic theory could relate directly. It thus seems to many that the search for substantive linguistic universals has been tried and has failed.

It is, however, entirely possible that many absolute and implicational universals are out there, but they can only be seen at a higher level of abstraction. Consider again the Verb–Object Constraint. When this is translated into a prediction about word order in the simplest possible way, we get a statement like "the verb and its object will be (nearly) adjacent to one another, whereas the verb and its subject need not be". This has the status of a strong statistical tendency (true for some 92% of languages with fixed word order) but not an absolute universal (false for some 8%). In other words, it is too promising to abandon, but there are certainly exceptions to it, just as Bickel says. Almost all of the exceptions are languages that have Verb–Subject–Object word order. Although noticeably much less common than Subject–Verb–Object order and Subject–Object–Verb order, it is not uncommon, and is found in different parts of the world. Closest at hand are the Celtic languages, which have been fairly well studied from a generative perspective:

(14) Gwelodd Siôn ddraig. (Welsh (Sproat 1985))
See.past.3sS John dragon
'John saw a dragon.'

Now there are several things to note in assessing this "counterexample" to the VOC. First, by all accounts, Verb–Subject–Object order is a relatively minor variant of Subject–Verb–Object order; only the position of the verb and the subject relative to each other is different. Adpositions still come before NPs, auxiliaries before main verbs, and so on. It is also notable that, in Welsh, SVO order surfaces when the tense marker and the verb do not fuse into a single word:

(15) Gwnaeth Siôn weld draig. (Sproat 1985)
do.Past.3sS John see dragon
'John saw a dragon.'

We also know independently that the verb moves to join with the tense marker in some languages and not others. For example, French and English are taken to differ from each other in just such a way. As a result, finite verbs come before adverbs in French ((16c)) but after adverbs in English ((16d)), even though nonfinite verbs come after adverbs in both languages ((16a–b)) (Pollock 1989).

(16) a. Jean a souvent embrassé Marie. (French)
b. John has often kissed Marie. (English)
c. Jean *embrasse souvent* Marie. (French)
d. John *often kisses* Marie. (English)

Second, there is some reasonably subtle evidence that the subject moves from the beginning of VP to the beginning of the clause in English. The evidence is that a quantifier associated with the subject can be "left behind" in the immediate preverbal position (Sportiche 1988):

(17) a. All the children will – find a candy.
b. The children will all – find a candy.

If we accept these two kinds of movement as legitimate grammatical processes, motivated to explain certain details about English and how it differs from French, then an easy account of Welsh emerges: Welsh is simply a language in which verbs move the way they do in French, but subjects do not move the way they do in English (Koopman and Sportiche 1991). So the possibility of Verb–Subject–Object languages emerges naturally out of possibilities that generative linguistics is committed to anyway. They are not exceptions to the VOC, when it is understood as an abstract claim about linguistic structure, not as a surface property of word order. And indeed, there is other evidence for the VOC in Celtic languages (Sproat 1985; McCloskey 1996).

Note also that a Verb–Subject–Object language is created only if a series of properties falls into place in a certain way: heads must come first in their phrases, verbs must raise to the tense position, and subjects must not move to the beginning of the clause. If each of these parameters of variation is set this way roughly 50% of the time, then we would expect to observe Verb–Subject–Object order in roughly 12.5% of the languages of the world—and that is not far from the observed figure. FGT can thus explain why one language type is less common than another when a constellation of factors is needed to produce that type (see Baker 2001, ch. 5 for more discussion).[6]

There are some obvious dangers here. It is certainly true that using abstract linguistic theories makes it harder to apply those theories to new languages. It also opens up opportunities for various kinds of cheating when it comes to evaluating hypotheses. For example, generative linguists might hide themselves from all counterexamples by making hasty appeals to movement, even when there is no evidence for it, the way that there is in Welsh. But do we have any reason to believe that we

[6] See also Chung 1998 for a different proposal for deriving Verb–Subject–Object order in Chamorro from a structure that obeys the VOC by way of a nontrivial movement process.

can do without it, even if it seems both safer and more convenient to do so? I think not, for at least two reasons.

First, when one studies one particular language in detail, one finds that a degree of abstractness is needed to give the best possible account of that language within its own terms. One familiar case in point is agreement paradigms: it is often necessary to posit morphologically null agreement affixes to complete a paradigm that has uniform grammatical behavior, or to distinguish two homophonous affixes that are actually agreeing with different underlying categories. Similarly, in the study of case inflections on noun phrases, one might have to posit a unified notion of (say) accusative case at an abstract level, even though it is realized by different observable morphemes in different circumstances (see, for example, Legate 2008, who shows how this is crucial to understanding the phenomenon of ergative case marking in some languages). Linguists may of course debate just how much abstraction is warranted when it comes to truly and fully describing the grammar of a particular language; I personally think that the answer is "quite a bit", especially as one tries to approach the generative ideal of a full description, one that does not presuppose "linguistic common sense" on the part of the grammar user but that tries to explicate what that "common sense" consists of.[7] But putting this debatable matter partly aside, suppose that we accept in principle that the grammars of natural languages studied in their own terms are found to be abstract to *some* nontrivial degree. Then it inevitably follows that comparing grammars to see how they are the same and different will have to be done at this same level of abstractness. There is no reason to think that one could avoid this abstractness except perhaps in a few lucky, nonrepresentative cases in which the abstract categories happen to map straightforwardly onto surface categories.

A second way of making essentially the same point comes from thinking again about the logic of Chomsky's acquisition-based sense of Universal Grammar. What kinds of knowledge of language does the generative linguist most want to attribute to the initial state of the human child? In point of fact, we typically want to attribute those facts that are *more* abstract and remote from obvious everyday experience to that initial endowment. There is (maybe) no serious language acquisition problem for the more obvious facts about word order, case marking, agreement, and use of lexical categories in a language. These matters are saliently and abundantly attested in the sample of language that is presented to any child, so they could be learned from the data by some kind of inductive process. The real acquisition puzzles come from those subtle and surprising but robust and replicable discoveries at the corners of a grammatical system that are often discovered by a formal-generative attention to detail, explicitness, and the special issues that can arise when simple structures

[7] Note that this scientific ideal is much more ambitious than what normally counts as "describing a language" in the current linguistic scene. Therefore, the fact that one might be able to do without significant abstractness in completing what now counts as a decent descriptive grammar of a language does not at all imply that abstractness is unnecessary for the ultimate goal of language description.

are composed to form more complex structures. These include distinctively generative discoveries such as the Binding Conditions on pronouns and anaphors, the so-called Island Conditions that constrain movement operations, conditions on the interpretation of quantified expressions, and so on. Often distinctive facts about these matters show up clearly and unambiguously only in sentences of a certain complexity, sentences with a very particular combination of properties. For such properties, it is hard to imagine that every language user that demonstrates the knowledge was exposed to a sufficient number of the relevant structures—*and* that they noticed them, *and* that they realized their significance for the grammatical point in question. These then are the grammatical properties that we have the most reason to attribute to Universal Grammar qua the innate endowment for language. It then follows that these are the grammatical properties that we have the most reason to expect to be universal also in the Greenbergian sense of being observable in all human languages. So we expect that the most abstract properties of language—the very hardest properties of that language for a linguist to discover, and thus the issues most rarely discussed in descriptive grammars—also to be the most universal properties of language.

Informal experience suggests that this may well be true. Of the ten or so non-Indo-European languages from various families that I have done serious fieldwork on, every one has phenomena that are recognizably like Chomsky's (1981) Binding Conditions and Ross's (1967) Island Conditions. (18) shows some familiar-looking contrasts from Mohawk, a language that is otherwise very different from English (Baker 1996, ch. 2); note that the Mohawk examples have the same grammatical status as their English translations.

(18) a. Úhka í-hs-ehr-e' Uwári ruwa-núhwe'-s?
who Ø-2sS-think-IMPF Mary FsS/MsO-like-HAB
'Who do you think Mary likes?'

b. *Ka nikáyʌ áthere' she-yʌtéri ne yakó-hs-u.
which basket 2sS/FsO-know.STAT NE FsO-finish-STAT
'Which basket do you know the woman who made (it)?'

c. *Úhka wa'-te-sa-hʌ'reht-e' ne tsi Uwári
who FACT-dup-2sO-shout-PUNC because Mary
wa-huwa-rasʌ'tho-'?
FACT-FsS/MsO-kick-PUNC
'Who did you shout because Mary kicked (him)?'

It is easy to multiply such examples, and no one is surprised these days when such things are discovered in a new language. But these matters are not considered at all in the *World Atlas of Language Structures*, nor in the sorts of typological databases that Bickel refers to. Of the 140 maps in *WALS*, not a single one concerns anaphor binding, quantifier scope, extraction from a complex noun phrase—or even the Verb–Object Constraint. They all have to do with word order, agreement, case marking, category membership, and other superficial morphological categories. Thus, standard typologists have looked hardest for universals in exactly those domains where generativists least expect to find them, and have hardly looked at all in those domains where generativists predict that they exist. It does not come as a surprise, then, that functionalist typology by itself has found little in the way of linguistic universals.

Why hasn't anyone done a *WALS*-style map that documents the distribution of (say) island effects in languages of the world? Even if one wanted to do so, it would be a complicated endeavor. While it is true that all the languages I have worked on have recognizable island effects, they are not all the same in this regard. These matters necessarily have to do with the interactions among phenomena. Therefore, any differences in the phenomena themselves will inevitably cause differences in the interactions that one would expect. It matters to island effects whether the phrase being extracted from sits in the object position or has been extraposed to the edge of a domain. It matters whether the extracted phrase is moved overtly or covertly. It matters whether the original position is occupied by a gap or by a resumptive pronoun. And so on. So even if a grammatical condition is universal in the strongest and simplest possible way, its observable effects will not be universal and invariant. For example, one can extract a question word out of the understood direct object in English but not in Mohawk:

(19) *?U'hka se-núhwe'-s ne ako-kára'?
who 2sS-like-HAB NE FsP-story
'Who do you like stories about?' (OK in English)

I believe that this difference does not undermine at all the claim that the same island conditions apply in both languages; on the contrary, the difference is expected once one realizes that the "object" is in a different position in Mohawk than it is in English (see next section). The point is that the observed effects inevitably depend on how the condition under investigation interacts with other aspects of the language, which may themselves vary. So the observables will be complex and multidimensional; it is not clear that they could be expressed in a useful and meaningful way in a simple map format, even once we pull together enough relevant material.

It is possible that the generativist's impression that there is something universal about the core island conditions (for example) will turn out to be an illusion. Perhaps the various language-specific island-like phenomena we observe cannot truly be unified into a single condition, and it is only because of vagueness in our understanding of the phenomena that we can see the conditions at work in different languages as being the same. It has happened in the history of generative grammar that some attempts to unify similar-looking conditions into a single, more general condition have either failed or proved to be Pyrrhic victories—claimed successes that were so complex or inflexible that they turned out to be the equivalent of a failure in the end. Much depends on whether the differences between the island effects observed in language A and those observed in language B can be correctly and insightfully attributed to independently observable differences in the phrase structure, or in the word order, or in what exactly is moving, or in where it is moving to, or in what is left behind by the movement (or something else). There are many encouraging results of this sort, and also many outstanding problems. But this discussion is enough to show why the putative result of functionalist typology that there are no solid linguistic universals does not really speak to the issue as it arises within a generative perspective.

12.3.4 Interim summary

Part of the motivation, then, for a distinct formal generative typology comes from the belief that standard functionalist typology is inadequate as a way of fully answering questions (1)–(3) because of its deep-set aversion to abstractness in linguistic analysis. Its techniques are not even adequate to tell us if something relatively straightforward like the Verb–Object Constraint is true or not. Many typologists have been guilty of a degree of laziness in not striving to understand the grammatical structures of the languages they are drawing on in their surveys, with the result that they may not ask the most important questions, cannot always filter out interfering factors, and are not in a position to recognize indirect and unforeseen consequences of the factor they are interested in. For the most part, they have not found the level of abstraction at which questions (1)–(3) can truly, insightfully, and productively be answered.

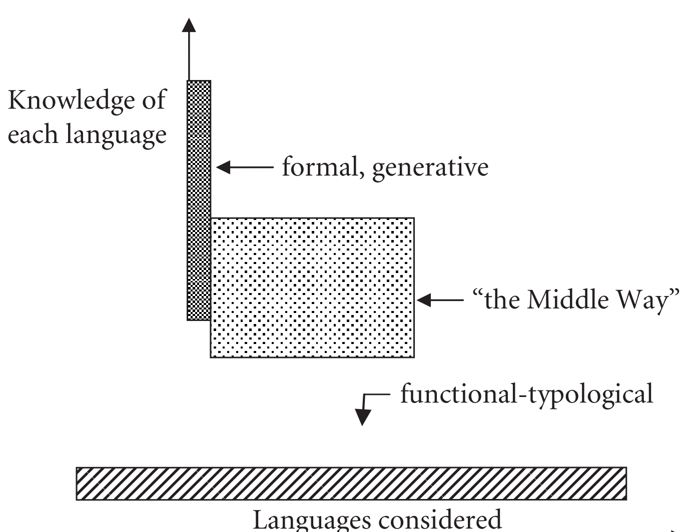
The other impetus for a distinct formal generative typology is that formal-generative linguists have also been guilty of a degree of laziness. We have usually not bothered to do the work we need to do to prove the genuine universality of our claims about Universal Grammar. It is all very well to predict that the surprising details discovered in the corners of one language will be universal, but we need to have methods for testing and refining these predictions. Despite the healthy sense that crosslinguistic comparison plays a larger role in generative work now than before, it still falls far short of what is attractive and right about the typological ideal. It is now commonplace to find examples from four or five languages cited in a generative linguistics article, but typically those languages are all from a single language family (usually Indo-European), or at most from two language families (for example, comparing an East Asian language to English). So while crosslinguistic comparison is on the rise, true typology is not. Conspicuously absent in most of this work is the typologist's vision of controlling for genetic and areal factors by sampling from unrelated languages and language families, and looking for naturally occurring replications of particular linguistic systems in other parts of the world. As a result, generative linguistics also has not made that much progress in establishing whether something like Verb–Object Constraint is universal or not. Generative linguistics has something important to learn from typological approaches in this regard.

12.4 A distinctive method

Of course, we are all lazy. More charitably put, we all have limited time, energy, and resources. Therefore, we need some realism as well as some idealism in our research enterprises. Although it would be great to know everything about everything, it is not feasible in this life, and we need to make choices. We also need to identify intermediate stopping points along the way where we can evaluate and take stock of our progress. How can FGT address this need?

My primary suggestion is to aim for what Baker and McCloskey (2007) dub "The Middle Way". This is simply to do an intermediate amount of linguistic research on an intermediate number of languages. There was a joke at the engineering school I attended that our school's strategy was to teach us "more and more about less and less until we knew everything about nothing". This was contrasted with the strategy of the liberal arts school up the street, which was said to be to teach "less and less about more and more until the students knew nothing about everything". The point of the joke is, of course, that despite the differing educational approaches, both student bodies end up in the limit knowing exactly nothing. Stereotypical formal-generative research risks achieving irrelevance via the tech school route, whereas stereotypical typological research risks achieving irrelevance via the liberal arts school route. The obvious alternative is simply to try to know something about something—the Middle Way. It is a simple mathematical fact that the way to maximize the area of a rectangle given a fixed perimeter is to make its height and its breadth equal. In the same way, linguistic understanding given finite resources is likely to be maximized by striking a careful balance between the range of languages considered and the depth of knowledge about each language, as sketched in (20).

(20)



In practice, this might involve working with samples of 10–20 unrelated languages, rather than with sample sizes one or two orders of magnitude greater, as is now common in typological studies. Each language, however, would be considered in generative-inspired detail, looked at in terms of its underlying grammar, not just its easily observed surface properties. Surprisingly few linguists are attempting this as a consciously-chosen strategy to identify universals and recurring patterns in diversity.

To illustrate the advantages of knowing an intermediate number of languages in an intermediate level of detail, consider evaluating the relevance of the Verb–Object Constraint for the Mohawk language. On first glance, this language is perhaps an even stiffer challenge to the VOC than Verb–Subject–Object languages like Welsh are. It is a language with no identifiable basic order at all (Mithun 1987), in which all of the following word orders are attested:

(21)	a.	Sak ranuhwe's ne atya'tawi.	(Sak likes the dress.)
	b.	Ranuhwe's ne atya'tawi ne Sak.	(Likes the dress Sak.)
	c.	Ranuhwe's ne Sak ne atya'tawi.	(Likes Sak the dress.)
	d.	Sak atya'tawi ranuhwe's.	(Sak the dress likes.)
	e.	Atya'tawi Sak ranuhwe's	(The dress Sak likes.)
	f.	Atya'tawi ranuhwe's ne Sak.	(The dress likes Sak.)

Hence no one well-defined and independently motivated kind of movement will save the day for the VOC, the way it did in Welsh. So the VOC might seem like a nonstarter for Mohawk.

But someone with a broader knowledge of Mohawk is in a position to see that this conclusion would be hasty. The VOC may not be visible in the *syntax* of Mohawk but it is visible in the morphology of Mohawk—in particular, in the phenomenon of noun incorporation. The object of the verb can optionally be realized inside the inflected verb as a kind of noun+verb compound ((22b)), but the subject cannot be ((22c)).

(22) a. O*wir*a'a wahrake' ne o'wahru.
baby ate the meat
'The baby ate some meat.'

b. Owira'a waha'wahrake'.
baby meat-ate
'The baby ate some meat.'

c. *O'wahru wa'ka*wir*ake.
meat baby-ate
'The meat was baby-eaten.'

Some other illustrative examples of object (but not subject) incorporation in Mohawk are given in (23).

(23) a. Wa'eksohare' 'She dish-washed.'
b. Wa'kenaktahninu' 'I bed-bought.'
c. Wa'khwistatshvri' 'I money-found.'
d. Wahana'tarakwetare' 'He bread-cut.'

Why does this difference exist? Although the details of different accounts vary, it is very plausibly related to the Verb–Object Constraint in (9). Indeed, it follows immediately from this constraint if we now interpret the "linguistic unit" that can contain the object and the verb but not the subject as being the inflected verb, rather than the verb phrase, a morphological unit rather than a syntactic one. We then observe the same constraint applying to different but analogous linguistic units to explain a significant parallelism in the data.

Our confidence that this is a related phenomenon is increased by considering English compounding. An understood object can also appear compounded with a deverbal noun in English, whereas an understood subject cannot. The examples in (24) and (25) are thus parallel to the Mohawk examples in (22) and (23) in this respect.

(24) a. meat-eating is disgusting. (= the eating of meat)
b. #baby-eating is disgusting.
(*not* eating *by* babies is disgusting; only the eating *of* babies)

(25) a. meat-eating, meat-eater
b. dishwashing, dishwasher
c. cheese-cutting, cheese-cutter
d. car-buying, car-buyer

So English compounding serves as a conceptual bridge. It is easy to see noun incorporation in Mohawk and compounding in English as two manifestations of a common underlying truth. It also not hard to see compounding in English and phrase structure in English as being related phenomena—a generalization first captured (for English) as Selkirk's (1982: 37–8) "First Order Projection Condition". Therefore, by transitivity, all three are related, and the Verb–Object Constraint applies to Mohawk as well as to English. This shows that, when doing typological investigation, we need to know a good percentage of the languages in our typologies well enough to recognize when there are nonobvious ramifications of potentially universal conditions or properties, or we may conclude that there are fewer genuine universals than there are.

We can also illustrate the advantages of the Middle Way from the negative side. Would more knowledge about Mohawk allow us to say anything about why Mohawk *sentences* do not show evidence of the VOC, even though Mohawk *verbs* do? Here I can give only the briefest sketch of a form of explanation that I worked out in detail in Baker (1996). The following sentences in English show that, although true objects must be next to the verb in English, "dislocated objects" need not be:

(26) a. That dress, Sak really likes it
b. Sak really likes it, that dress.

We would not say that a sentence like (26a) violates the VOC: the object for grammatical purposes is the pronoun, and this pronoun does form a phrase with the verb. *That dress* is not the direct object but a topicalized phrase that the direct object refers to. Now Mohawk could also have object dislocation, and that would not tell against the VOC. But what would dislocation look like in Mohawk? It so happens that in Mohawk, weak subject and object pronouns are phonologically null. Thus, a sentence like "He likes it" can be expressed in Mohawk just by the inflected verb.

(27) Ra-nuhwe'-s.
MsS/NsO-like-HAB
'He likes it.'

Nor is it surprising that Mohawk allows null pronouns. Mohawk verbs bear prefixes that agree with both the subject and the object; for example, the prefix *ra-* in (27) indicates that the subject is masculine singular third person and the object is neuter third person. Because this information is expressed in the verbal morphology, overt pronouns are not needed, just as subject pronouns are needed in English but not in Spanish or Italian.

Now, given this simple fact about Mohawk, what would dislocation sentences analogous to (26) look like in Mohawk? They would look like (28).

(28) a. Atya'tawi Sak ranuhwe's.
dress Sak likes (it)
b. Sak ranuhwe's ne atya'tawi.
Sak likes (it) the dress

In fact, they would look like the grammatical sentences in (21a) and (21e). We thus have the appearance of free word order of the object, and with it violations of the VOC. But in English we say that these are not real violations of the VOC, because the true object is the pronoun, in the verb phrase, adjacent to the verb, where it belongs. We can say exactly the same thing about Mohawk. (28a)/(21e) is not a counterexample to the VOC, either: the object is in the verb phrase, next to the verb, in Mohawk, too, for all we know—we just don't see it because object pronouns are null in Mohawk, for reasons that are both predictable and independently observable.

Is this just a slick trick, or is there other evidence that sentences like (28) are more like dislocations in English than they are like simple (6) in English? In Baker (1996, ch. 2) I argued at length that there is lots of independent evidence for this. One class of evidence comes from nonreferential NPs, such as anaphors, nonreferential quantifiers, and idiom chunks. We know that in well-studied languages like English and Italian these sorts of nonreferential NPs cannot be dislocated (**Nobody*/**himself,* *John saw him in town*; Cinque 1990). Furthermore, Mohawk does not have independent NPs of these sorts; there is no NP that is an exact equivalent of *nobody* or *herself* in Mohawk, for example. This is just what we expect if overt "objects" are actually dislocated adjuncts in Mohawk.

In this light, recall from section 12.3.3 that one cannot extract a question word from a putative direct object in Mohawk the way one can in English (see (19)). Now it so happens that one cannot extract a question word out of a dislocated object in English either, as shown by the contrast in (29).

(29) a. Who did you hear a story about?
b. *Who did you hear it, a story about?

The contrast is expected: a dislocated object is really a kind of adjunct, so extracting from it is a kind of adjunct island violation, bad for the same reason that the English version of (18c) is. If apparent objects in Mohawk really have the status of dislocated NPs that are indirectly related to (null) pronouns which are the true objects, then we would expect that one could never extract from these apparent objects in Mohawk—especially since we already have some evidence that the adjunct island condition holds in Mohawk as well as in English ((18c)). Thus, we actually predict the badness of (19). At the time, this seemed like an anomalous fact, calling into question the universality of the island conditions. But now it does not seem anomalous at all; rather it is expected given what else we know about the language.

So the story holds together. Once we understand the grammar of Mohawk to some degree, we realize that we do *not* expect to observe the VOC in Mohawk syntax, because its agreement properties mean that object dislocation is always a possibility and will not look obviously different from nondislocation on the surface. So some real knowledge of the grammar of the language is necessary to evaluating the universality of a condition like the VOC (or the adjunct island condition), not only so that we can recognize reflections of the VOC that we might not have thought to look for but also so that we do not look for simplistic evidence of the VOC in places that we really should not expect it. Practicing the Middle Way makes this feasible.

Will the results of research done in these ways scale up to larger, more impressive sample sizes? There is no guarantee, of course; this is research, and once we know for sure exactly how to do it, that phase of research is probably almost over. But my recent investigation of the principles and parameters of agreement systems (Baker 2008*a*) suggests that the answer is positive. In this study, I began with a close comparison of agreement in some Bantu languages with agreement in Indo-European languages, then moved to a pilot study of 10–15 languages (the Middle Way stage), and then tested my two parameters against a larger sample of 108 languages, based on the core languages survey of *WALS*. I found that the move from 15 languages to 108 did not give appreciably different results, although it did of course serve to illustrate those results in a much wider range of language types. (For example, one could see what different agreement systems look like within languages with different word orders, including unusual ones like Object–Verb–Subject and Object–Subject–Verb.) My guess is that one does not have to look at a very large number of unrelated languages to get most of the benefit of guarding against parochialism in one's linguistic analysis. Nor does it take that long to learn what one needs to know about a language in order to guard against superficial errors and misinterpretations. Along both dimensions, the famous 80–20 rule that 80% of the results are gained from 20% of the effort probably applies. If so, "Middle Way" methodology should be very effective.

Saying that this is what needs to be done by the field as a whole does not imply that this is what each linguist needs to do. There are obvious ways to contribute to an FGT-style program by attempting only pieces of it. For example, doing a generative analysis of a single understudied non-European language that discovers the right level of abstraction for capturing facts about that language, using data and terminology that are accessible to others, is a huge contribution. And linguists can work in teams, use each other's results, and corroborate each other's finding to accomplish a larger portion of this than any of them could do by themselves. In a variety of such ways, one can realistically hope to learn something about something.

12.5 On the relationship between the subfields of linguistics

What does FGT have to say about the relationships between the different subfields of linguistics? Since it is more of a research methodology than an independent theory, it has no fixed distinctive position on those matters. What answers it has are either taken over from its Chomskyan/generative inheritance or are discoveries it claims to have made in the course of its pursuits.

FGT certainly assumes a lexicon, which is at least a list of the atoms that can be used in a syntactic representation, together with the properties that distinguish one from another. It is quite possible that the lexicon is also no more than this (Marantz 1997). Syntax, then, is the system by which such elements can be combined into larger, well-formed and coherent linguistic representations. Phonology and semantics are both taken to be interpretive, in the usual Chomskyan sense. In other words, we assume for theoretical purposes that a formal syntactic representation is constructed, and then phonological principles calculate how to pronounce that syntactic representation and semantic principles decide what it could be used to mean in a particular context (see also below).[8]

12.5.1 Morphology and syntax

The place of morphology is a particularly interesting question within this approach. The layout just sketched leaves it as an open empirical question whether the atoms listed in the lexicon correspond more closely to the traditional idea of a word, or to the traditional idea of a morpheme. In fact, there turn out to be many similarities between word structure and sentence structure, such that we can capture significant generalizations by having abstract principles that are relevant to the formation of both. The Verb–Object Constraint is an excellent case in point. The theme/object NP but not the agent/subject NP combines with the verb to make a verb phrase in English. Similarly, a noun expressing the theme/object but not a noun expressing the agent/subject can compound with the verb root to make a complex verbal stem in Mohawk. We want to say that this is not an accidental similarity; rather, the two are both reflections of the same underlying fact about human language. This is most naturally done if we do not consider morphology and syntax to be two separate domains but rather include both in the same comprehensive representational system, so that both are subject to the same representational laws.

In fact, there are many ways in which what shows up as a constraint on the syntax in a language like English shows up as a constraint on word structure in Mohawk and similar languages. A second example is the so-called Extended Projection Principle which stipulates that clauses must have subjects, but makes no such requirement about objects (Chomsky 1981). This can be seen in English syntax in the fact that meteorological predicates like *rain* must have a dummy placeholder pronoun in the subject position but not in the object position:

(30) a. *Rained yesterday.
b. It rained yesterday.
c. *Rained it yesterday.
d. *It rained it yesterday.

This constraint is not readily observable in the syntax of Mohawk or Mapudungun, since these languages omit all unstressed pronouns. But it can be seen in the morphology of Mohawk and Mapudungun. These languages require the verb to

[8] Exactly what sequence of calculations an actual language user might go through in accomplishing a particular task that uses language is taken to be a partially different matter, and is left largely open. This is a matter of *using* a language rather than a matter of *knowing* a language, and understanding what it is to know a language is assumed to be a necessary although not sufficient condition to understanding what it is to use a language.

bear a subject agreement affix, but they do not require the verb to bear an object agreement affix:

(31) a. v-yo-kvnor-e' (*v-kvnor-e') (Mohawk)
FUT-NsS-rain-PUNC FUT-rain-PUNC
'It will rain.'

b. Petu mawün-üy. (*mawün-fi-y) (Mapudungun)
ADV rain-IND.3sS rain-3sO-IND.3sS
'It is raining.'

Yet another example of this concerns anaphora. English contains a difference between the overt pronouns *him* and *himself*, such that the marked pronoun *himself* must be used if and only if a reflexive interpretation is intended (Reinhart and Reuland 1993). Mohawk does not have a difference between two pronominal forms; there is only one masculine singular pronoun, *rauha*. But Mohawk does have a parallel morphological difference: a verb marked by the special prefix -*atat*- must be used if and only if a reflexive interpretation is intended.

In all three cases, the same (abstract!) constraint seems to be at work, even though its effects show up most clearly in the syntax of some languages and in the morphology of others. Hence, typological research that is open to abstract generalizations discovers that morphology and syntax are subject to the same potentially universal principles. This in turn suggests that they are not fundamentally different areas of study but aspects of the same grand compositional system. (See also Cinque 1999 for ordering conditions that govern both the positioning of adverbs in some languages and the order of tense–mood–aspect morphemes in other languages.)

12.5.2 Syntax and pragmatics

The results of FGT also have implications for controversial questions about the relationship between syntax and pragmatics. Everyone agrees that the two are interrelated in some way. The controversial question is which one leads and which one follows. On one view, pragmatics is the more basic study, and syntax is the crystallization (grammaticization) of pragmatic functions into more or less iconic grammatical forms. On the other view, syntactic principles determine what sentences can be formed, and then pragmatics takes the range of syntactic structures that are possible and assigns to each of them some natural pragmatic use(s) that take advantage of the grammatical forms that are available. The first view is characteristic of functionalist approaches to linguistics; the second is the traditional Chomskyan position.

What we have seen about the structure of Mohawk seems relevant to this matter. English permits both structures in which the direct object is in place and structures in which the direct object is dislocated. These two structures have distinct pragmatic uses, in a consistent, natural, iconic way. In contrast, Mohawk permits only the dislocation structure, (in part) because agreement with the object is grammatically necessary in Mohawk. Therefore, only one of the two structures is available. The dislocation structure is thus forced to cover a wider pragmatic range in Mohawk than in English. For example, only definite NPs with a topic reading are dislocated in English, but there is no such requirement on dislocation in Mohawk: indefinite, nontopical NPs are possible in the very same peripheral positions that definite NPs are in Mohawk.

There is another domain in which Mohawk allows a larger set of possibilities than English does: Mohawk allows an object to be incorporated into the verb, whereas English does not. As a result, Mohawk speakers can draw a pragmatic distinction between incorporated nouns and unincorporated ones: incorporated nominals are backgrounded, and contrastive focus is incompatible with incorporation. In contrast, ordinary independent objects are used for backgrounded as well as for focused interpretations in English, that being the only relevant structure available.

Therefore, it seems then that pragmatic considerations—which are taken to be more or less universal—cannot in themselves explain the different ranges of structures that are available in different languages. What pragmatic considerations can do is, given a range of well-formed structures, say something about what each one might naturally be used for. That is just what one would expect if semantics/pragmatics has an interpretative relationship to syntax, rather than a formative one. I am not in a position to say that this state of affairs cannot be made sense of from a functionalist, pragmatically-driven perspective; they are certainly well aware of such facts. But at the very least, there is no reason here to abandon the syntax-driven approach inherited from FGT's generative roots. On the contrary, FGT-style research seems to confirm that form is partially distinct from function.

12.6 Types and levels of explanation

This leads naturally to the question of what kind of explanation does FGT seek to offer, and what relationship does it see between language and other areas of cognition. The short answer is that it offers language-internal explanations, rather than language-external ones. One feature of a language is explained in terms of its similarity to another, at first different-seeming feature of that language or another language, by saying that both are consequences of the same general principle. As such, the explanation of a feature is *not* primarily in terms of its pragmatic value, or its iconic relationship to other cognitive systems, or in terms of its diachronic origins.

To take a familiar example, that also have object–verb order (see (11)) simply because NP and P constitute a phrase, the object and the verb also constitute a phrase (see (9)), and there is one general rule for how phrases are built in a language, with the "head" of a phrase either coming first (English) or last (Japanese), all other things being equal. These two sorts of phrases have common properties because they are both products of the same phrase-building machine, just as two bullets might have similar markings because they were both shot from the same gun.

This is not to deny that there may ultimately be external, functional explanations for some of these matters as well. But even if the language internal explanations of FGT are not ultimate explanations, they are still real, significant, plentiful, and valuable. Indeed, part of their value is that they clarify the full nature of natural language, with all its interconnections, and thus reveal what its deepest properties are that may call for another level of explanation. For example, the discussion above suggests that the Verb–Object Constraint might well be a universal property of human languages. If so, why should this be? I do not know, and I have never heard someone attempt a really serious explanation, even though the basic claim has been known to generativists for some time. It may very well show us something deep and important about how human beings conceive of events and the entities that take part in them, something that might have been difficult or impossible for us to discover otherwise.

Some might take it as a weakness of FGT that it offers no specific tools for answering "type (4)" questions like this. But what I find exciting about FGT is that it does succeed in *raising* such questions. The VOC could be a true universal of language, and it is certainly true of a very wide range of languages, even though it shows up in different ways in different languages. No other approach could really make and confirm this discovery: it is too abstract for functionalist typology to find, whereas it is too sweeping a claim for a nontypological generative approach to confirm. Much FGT style research has gone into this discovery, and much more may still be needed to see if it is really valid for all the languages of (say) Amazonia and New Guinea.[9] If in the long run FGT does nothing more than set the proper stage for a true Minimalism or a true eliminative functionalism, I think that it has done a lot of

[9] A full discussion of the VOC should talk about ergative languages, including the gradual rejection of Marantz's (1984) "deep ergative" hypothesis in favor of analyses that are compatible with VOC (for example, Bittner and Hale 1996). On the negative side, Warlpiri is an interesting case of a nonconfigurational language that has been fairly extensively studied, but has yielded little evidence for the VOC, at least in syntax. I for one remain uncertain just how this language fits in.

valuable work that it can be proud of. (I also think it is a virtue of the approach that it does not push one to offer hasty and premature pseudo-explanations, which may offer some rhetorical satisfaction but offer little in terms of discovery or prediction. But I do press this point here.)

12.7 Types of evidence

The sources of data for FGT research are simply the traditional ones of the field. For example, it sees the value of both targeted elicitation techniques and the study of naturally occurring data in recorded narratives or conversations. Each kind of data can be used to complete and correct the other. One can even go back and forth between the two. A naturally occurring example might be a great starting point, as a model for targeted elicitation. Conversely, it might help to ask a speaker to build a small narrative around a sentence that arises in elicitation. Any behavior that is a product of true linguistic knowledge and use is in principle welcome.

Large-scale corpus techniques are probably not especially useful in practice, simply because large corpora are not available in the range of languages that need to be considered. But you never know. The same point holds for psycholinguistic and neurolinguistic experimentation. My impression is that, in order to answer the explicitly comparative questions in (1)–(3), what we most need is first-pass generative-style analyses of a wider range of languages, rather than new techniques that can mainly be applied to more familiar and accessible languages. There are hard practical problems about how to do psycholinguistic and neurolinguistic tests on languages spoken in remote areas by elderly people of a very different culture, and these may just not be worth solving, for now. But I hasten to add that questions (1)–(3) are not all there is to linguistics. Other techniques may be needed to answer other important questions, and convergence across different modes of inquiry is always to be looked for and hoped for.

Finally, FGT makes no special use of sociolinguistic and diachronic data. That may partly be a deficiency in our training. Surely these areas of inquiry are related, at least in the sense that diachronic changes have produced the range of languages we now have to study, and sociolinguistic variation probably helps to create diachronic change. In answering the question of whether crosslinguistic differences are patterned or not (question 3), looking at which languages can develop from a given language type could be particularly valuable. For example, the fact that uniformly head-initial languages like English and French have evolved from the same source as uniformly head-final languages like Hindi helps to suggest that there is a unified word order parameter. Similarly, Baker (1996) takes the fact that uniformly polysynthetic languages have evolved out of nonpolysynthetic languages in Northern Australia as evidence that there is a polysynthesis parameter, and Baker (2007) draws similar morals about agreement from comparing Makhuwa with its Bantu cousins and Ecuadorian Quechua with Peruvian Quechua. We thus use some quasihistorical examples on an opportunistic basis, and there may be opportunities in the future to do this more deeply and systematically.

But there are limits to the potential of such inquiry, too, which keep it off the top of our (my) agenda. First, it is not always necessary to know where the diversity we see came from in order to answer questions (1)–(3) and the implications of those answers for human cognition more generally. Second, FGT accepts Chomsky's point that the most real and primary notion of language is the mental representations in the minds of individual speakers. Those typically do not represent the historical changes that led to the current language. In principle, they would be exactly the same even if very different paths of historical change led to the same primary linguistic data. In short, the history of a language is not known to the average native speaker of a language, so it can have at most an indirect role.[10] Added to this is the practical problem that we have very little recorded history for any of the languages spoken outside of Eurasia, and reconstruction of syntax by internal reconstruction and the comparative method seems like a rather speculative enterprise. We do know that there has been enough historical change to get linguistic diversity in different areas of the world; therefore the universals that we observe do not seem to be simply due to inheritance from a common ancestor. That might be most of what we really need to know to face problems (1)–(3).

12.8 Conclusion

Formal Generative Typology is a pragmatic and somewhat eclectic approach that is built on the idea of combining a generative-style toleration for abstractness in analysis with a typology-inspired interest in testing claims over a sample of unrelated languages. This seems to be a powerful combination, with the potential to greatly advance the quest for answers to questions about what is truly universal to natural human languages, what can vary, and whether the variation is patterned or not.

[10] In contrast, some sociolinguistic variables probably *are* represented in the minds of speakers to some extent, since speakers can use and understand different styles in different social contexts. Perhaps, then, these should have a larger role in FGT (and generative linguistics more generally) than they have to date.

Further Reading

An accessible, book-length overview of this approach is Baker (2001). More specialized methodological discussions are included in Baker and McCloskey (2007) and Baker 2008*b*. Some canonical examples of the approach in action include Baker (1996), Baker (2003), Baker (2008*a*), and Cinque (1999). For more information about the generative substrate that the approach is built on, see the references in Boeckx (this volume).

CHAPTER 13

A FRAMES APPROACH TO SEMANTIC ANALYSIS

CHARLES J. FILLMORE
COLLIN BAKER

In this chapter, we contrast a broad use of the term *frame* in cognitive science with its related use in a type of linguistic analysis; we describe the principles and data structure of a particular research project (FrameNet) as a model for representing frame-based analyses of lexical meanings; we briefly introduce an extension of the project to include the semantic contributions of grammatical constructions; and we conclude by surveying the implications of a frames perspective on some familiar issues in linguistic semantics.

13.1 Frames and frames

Any discussion of a "frames approach" to semantic analysis must first draw a distinction between (1) the ways people employ *cognitive frames* to interpret their experiences, independently of whether such experiences are delivered through language, and (2) Frame Semantics as the study of how, as a part of our knowledge of the language, we associate *linguistic forms* (words, fixed phrases, grammatical patterns) with the cognitive structures—the *frames*—which largely determine the process (and the result) of interpreting those forms.

13.1.1 Cognitive frames

There is a general concept of *frame* (Minsky 1975; 1988; Goffman 1974; Tannen 1993), together with allied concepts like *schema* (Bartlett 1932; Rumelhart 1975), *idealized cognitive model* (Lakoff 1987), *script* (Schank and Abelson 1977), and even *meme* (Dawkins 1976), *narrative*, etc.), especially as developed in the cognitive sciences since the 1970s, that can be defined as any of the many organized packages of knowledge, beliefs, and patterns of practice that shape and allow humans to make sense of their experiences. Frames, in this sense, play an important role in how people perceive, remember, and reason about their experiences, how they form assumptions about the background and possible concomitants of those experiences, and even how one's own life experiences can or should be enacted.

Cognitive frames are usually expressed as "slot-filler representations", structures of interconnected roles together with constraints on the possible or likely fillers of those roles (Brachman and Schmolze 1985). Examples of such frames are (1) the partially ordered set of events, as well as the participants in such events, that one can expect in a typical visit to a restaurant, barbershop, or hospital, (2) stages and processes in the life cycle of a human being, (3) the visual and physical properties of a cube, and (4) the organization of a human face, and countless others.

As humans we have access to some of these frames by virtue of living on the earth, subject to its daily and annual cycles and the entities that we perceive; other frames we owe to just being human, with bodies that respond to gravity and to our biological and emotional needs, and with the perceptual faculties that our bodies possess; others we have by being members of a particular culture, where we consciously or unconsciously respond to its institutions, symbols, artifacts, and values; and, importantly, still others we have by virtue of being a part of the specific speech community that supports and is supported by the culture. Thus, we have schematic knowledge about gravity, heat, and shadows, the difference between living and non-living things, about colors, pain, joy and jealousy, about marriage, government and religion, and about weekends, heartburn, military titles, the color purple, and bikinis.

As an example of how the availability of a cognitive frame can shape our perceptions, independently of language, imagine a simple visual experience. In an

American setting, when we see a group of neatly dressed children approaching someone's house, carrying brightly wrapped packages, we are likely to interpret this percept by summoning from our memory what can be called the birthday party frame, which leads us to infer that some child is celebrating a birthday, and to expect that the children's experiences during the time of the party will include games, toys, and cake, and lighted candles on the cake, as well as a singing of the birthday song.

The following text, distributed on the internet some years ago in joke mailing lists, can illustrate how a number of bits of frame knowledge can be assembled to fill in the details of a larger "composed" frame. The title of the passage is "A Woman's Dream Breakfast". It will be useful to consider its intended interpretation.

> She's seated at a table in her garden with her gourmet coffee. Her son's picture is on the Wheaties box. Her daughter's picture is on the cover of *Business Week*. Her lover's picture is on the cover of *Playgirl*. And her husband's picture is on the back of the milk carton.

Understanding this little vignette requires an appeal to a great many facts about current American culture.

(a) Having breakfast at a table in one's garden is a common image of luxury in Western culture, enhanced by the explicit mention of gourmet coffee. Our heroine leads a comfortable life.
(b) Wheaties (a wheat and bran breakfast cereal) with milk is a stereotypical American breakfast.
(c) The pictures on Wheaties boxes are of popular athletes. Our heroine's son has done well in the sports world.
(d) Having her picture on the cover of *Business Week* surely means that our heroine's daughter has been successful in the business world.
(e) The magazine *Playgirl* features good-looking well-built young men, and one of those, we learn, is her lover.
(f) Currently the backs of milk cartons often show appeals for help in finding missing people: this woman's husband has disappeared and may never be found. The title of this vignette, "A Woman's Dream Breakfast", suggests that having her husband out of sight is an important component of our heroine's happiness.

Its distribution in a series with other texts intended to be witty suggests that this text belongs to a genre of humor known as "the battle of the sexes". A full interpretation would include the meta-judgment that the joke is probably more amusing to women than to men.

Obviously someone who is not familiar with these bits of cultural knowledge will not share the intended associations: the frames we invoke are based on knowledge we have, not so much about the words, as about the phenomena and their association with cultural values. (That is why it is difficult to understand a joke in a foreign language or one that expresses an unfamiliar culture.)

The intended interpretation of this text, as reconstructed here, illustrates well the common assumption, among semanticists, that the full meaning of a text is "vastly underdetermined" by its linguistic form alone. A dictionary, no matter how detailed, will not help the reader of this text if it is limited to the "standing definitions" of its individual words; and we cannot easily imagine that a computer program processing this passage would be capable of recognizing it as humorous.

The earlier invocation of the birthday party frame was based on a visual experience; the invocation of the complex structure that constituted the interpretation of the dream breakfast depended on linguistically presented information, but it was just a series of existence statements. The birthday party frame was first introduced into the artificial intelligence literature by Marvin Minsky (1975) where it was offered as an explanation of the most natural interpretation, not of a visual experience, but of a very short English text:

Mary was invited to Jack's party. She wondered if he would like a kite.

This time the explanation involves the evocation of a group of frames by particular linguistic features, but it is completed by the knowledge, on the part of an American reader, of the details of a child's birthday party. The linguistically anchored frames are evoked by the verb *invite*, which designates a relationship involving a host, a guest, and an occasion, the noun *party*, evoking a social event that will generally have a host, guests, and an occasion. The genitive construction *Jack's party* could either be interpreted as a party for which Jack is the host or one for which Jack is being celebrated. There are no linguistic traces of the birthday party frame as such, but the noun *kite* designates a child's toy, and the concern is whether Jack would *like* one. These various language-provided details jointly allow the reader to summon up a frame into which they could be articulated, where Jack is the birthday celebrant, the kite is being considered as a birthday present, and so on.

In cases like the "Dream Breakfast" vignette we say that the interpreter *invokes* the (cognitive) frames that enable the experience to make sense, whereas in the cases we will be concerned with in the rest of this chapter we will say that a given linguistic sign *evokes* the (linguistically anchored) frame that contributes to interpreting the passage. The birthday party text provided information to which the invocation of the birthday party frame assigned coherence. Frame invocation is a cognitive act that the interpreter (possibly quite unconsciously) performs to make sense of some incoming information. By contrast, frame evocation is a cognitive experience on the part of an interpreter that comes about by the interpreter's responding to language-specific associations connecting linguistic signs with particular frames. The discovery and analysis of such associations, those acquired as a part of learning the language, constitute Frame Semantics.

13.1.2 Frame Semantics

We take *semantics* to be the relation between linguistic forms and their meanings (distinct from semantics as a scheme for meaning representations). Thus, semantic analysis proper is the process or activity of showing how features of meaning are anchored in features of linguistic form. The language analyst can see this relationship from either direction. It has a *decoding* or *semasiological* aspect, by which the semantic import of linguistic structures is explained, and an *encoding* or *onomasiological* aspect, by which concepts are examined with respect to the ways in which they can be linguistically expressed. In Frame Semantics in particular, the meaning dimension is expressed in terms of the cognitive structures (frames) that shape speakers' understanding of linguistic expressions.

One part of language learning is acquiring the linguistic coding of already familiar experiences. For example, surely children are familiar with experiences of pain before they encounter the various linguistic means of talking about them. These include utterances expressing a current experience of pain (*ouch!*), clauses associating pain with the body, parts, or whole (*my foot hurts, my tooth aches, I have a pain in my foot, I hurt all over*), clauses associating pain with particular causes of pain (*the injection won't hurt, these shoes hurt, you're hurting me*), and ultimately an abstraction designating the quality of such experiences (*pain is a powerful motivating force*). In effect, dictionary definitions of the words like *pain*, *hurt*, *ache*, and *ouch* cannot (and need not) "explain" these concepts, but can only give pointers to the kinds of experiences with which the learner is already familiar.

A second part of language learning consists in acquiring new concepts—new frames—together with the linguistic means of coding them. In many cases a lengthy chain of prior learning is a prerequisite to attaining the new concept, as with the mathematical concept *mantissa*, which requires previous familiarity with such concepts as base, power, logarithm, decimal point, and, of course, the conceptual prerequisites of each of these in turn.

Thus Frame Semantics is the study of how linguistic forms *evoke* or activate frame knowledge, and how the frames thus activated can be integrated into an understanding of the passages that contain these forms. The full study of the understanding process, as seen in the Minsky text, must also take into account the ways in which non-linguistic information is integrated into the process.

With respect to the lexicon, we say that each *lexical unit*, the pairing of a word with one of its meanings (Cruse 1986), *evokes* a frame and *profiles*[1] some aspect or component of that frame. The pairing of a word with its background frame means

[1] R. Langacker (1987) provided the parade example of the pairing of the meaning of a word and its background in the concept of the *hypotenuse*: no definition of hypotenuse can be successful without making clear the associated notion of the right angle triangle, since a hypotenuse is the slanted line in just that kind of geometric figure. Langacker contrasts the *base*, corresponding to the background frame, and the *profile*, identifying the concept that requires that background. The usage above has borrowed *profile* as a verb.

that when we understand a word, we must simultaneously recognize the relevance of the background information within which that word plays an interpretive role (Fillmore 2003). In the case of specialist language, this bipartite nature of lexical knowledge is familiar from common lexicographic practice. For example, dictionary entries for *id*, *ego*, *superego*, *transference*, etc., in the psychoanalytic senses, are likely to indicate the connection to the theories of Freud and his followers: only by understanding the basic outlines of the background theory do these words make any sense. Similarly, a basic understanding of tectonic plate theory in geology is necessary in order to make sense of terms like *subduction zone*, *transform boundaries*, or *continental plate*. In most cases, for obvious reasons of space, modern dictionaries will only include a pointer to the needed domain, and it is up to the user to seek out the needed information.

The basic assumption of Frame Semantics, in respect to word meanings, is that not just words from technical domains but essentially *all* content words[2] require for their understanding an appeal to the background frames within which the meaning they convey is motivated and interpreted.

The background knowledge assigned to frames is often so thoroughly "overlearned" that considerable cognitive effort is sometimes required to bring it to consciousness. For example, we cannot understand the meaning of *Tuesday* without knowing about how time is reckoned in Western culture, including the established cycle of seven days and the naming of the members of the cycle. The concepts *weekday* and *weekend* depend on still more knowledge, this time involving designated members of the cycle typically dedicated to work and non-work. Understanding an expression like *Thank God it's Friday!* depends in turn on that distinction and an assumed natural preference of non-work over work. Similar principles apply to categories such as the four seasons, the compass directions, *up* vs. *down*, *right* vs. *left*, and thousands of others. These lexical units all call on shared background conceptualizations and are best learned, and understood, in packages, large or small.

Much of the early informal work in Frame Semantics offered descriptions of individual words, or small lexical systems linked by simple contrasts, that required appeal to background motivating contexts. Some of these backgrounds can be of considerable complexity, as in examples like the following:

Alimony. *Leo has missed three successive alimony payments.*

To understand *alimony* requires understanding divorce and the kinds of contracts that can be entered into at the time of a divorce; to understand *divorce*

[2] That is, nouns, most verbs, adjectives, demonstratives, adverbs, many prepositions, etc. The *function* words (articles, complementizers, prepositions, support verbs, etc.) contribute to meanings only as components of particular grammatical constructions. Theories differ as to whether it is useful to think of such words as contributing to the meanings of the structures that contain them.

requires knowing about marriage and its commitments; to understand *missing three successive payments* requires knowledge of the kinds of periodic commitments undertaken with an alimony agreement, and to know that missing a payment is a failure to honor such a commitment. Describing the payments as having been *missed*, rather than using a neutral expression like *did not make* (the payments), reveals an evaluation that this is a failure on Leo's part. All of this comes with knowledge of meanings of the associated words, in this case centered in the concept of *alimony*.

On land. *The men were happy to spend several hours on land this morning.*

The first definition of the noun *land* in the Concise Oxford Dictionary (COD New Edition) is "the solid part of the earth's surface"; the entry refers to its opposition to *sea*, *water*, and *air*. The prepositional phrase *on land*, however, is specifically understood as being in a contrast set with the phrase *at sea*, indicating location in respect to a water mass; learning that someone has spent a limited amount of time *on land* invites the inference that these several hours constituted an interruption of a sea voyage. The pair *on land* vs. *at sea* is matched for different contrasts by such pairs as *on the ground* vs. *in the air*, *on earth* vs. *in heaven*, *on earth* vs. *in (outer) space*. (See Fillmore 1982; 1985)

Pedestrian. *Try not to hit any pedestrians.*

The noun *pedestrian* refers to someone walking in an area where there is competition for space between moving vehicles and persons on foot, and so the concept of *hitting a pedestrian* must be understood from the point of view of a vehicle and its drivers: what is to be avoided is allowing one's vehicle, not one's fist, to strike a walking person. The word, furthermore, does not designate a member of a category of persons, but a person in a currently-relevant role. The frame that comes with *pedestrian* immediately enables an envisionment of the context for such an utterance; no such rich envisionment would be linguistically provided for an utterance like *Try not to hit any Norwegians.*

Decedent. *The decedent lived most of his life abroad.*

The word *decedent* is defined in COD as "a deceased person"; *deceased* as an adjective is defined as "dead" and as a noun it is defined as "a person who has died, esp. recently" ("especially recently"!). It seems strange that the nouns *deceased* and *decedent*, always used with a definite determiner, referring to some specific now-dead individual, should be defined with an indefinite NP. The sentence, of course, cannot mean that this person was dead when he lived abroad, but only that the individual referred to, now dead, lived abroad when he was alive. The "framing" of *decedent* as opposed to *deceased* is more complicated still: it is a term of art in U.S. law, and it is typically spoken or written only in discourse about the dead person's estate. The most common co-occurring words in sentences containing *decedent* are *probate*, *estate*, *heir*, *taxes*, and the like.

With nouns that designate an object that has a function, the linguistic context can sometimes distinguish between cases when it refers merely to the physical object and cases when it presents that object as serving its function: Pustejovsky (1995) distinguishes these as the *formal* and the *telic* properties of a noun. If someone is seated inside a bus, one can truthfully describe that person as being *in the* bus, even if the vehicle is a long-abandoned bus with its wheels missing. If we are told that the individual is *on the bus*, however, more assumptions are needed, related to our understanding of what a bus is for: in particular we assume the bus is "in service". A similar association with the telic of a noun holds for the contrast between being *in jail* vs. *in the jail*. We assume that someone who is *in jail* is being punished for some crime; someone could conceivably be *in the jail* to get out of the rain; and analogous contrasts hold for being *at school* vs. *at the school, in church* vs. *in the church*.

Since the ground observations about Frame Semantics must be the ways in which users of the language understand what is communicated by their language, Frame Semantic research is necessarily *empirical, cognitive, and ethnographic* in nature. Researchers must *find out* what frames inform the language being studied because there is no place to look it up; it involves subtle issues of language understanding rather than symbol manipulation and simple judgments of truth; and it requires learning about the experiences and values in the surrounding culture.

13.2 Frame Semantics and the FrameNet Project

The examples considered so far have all been accounts of individual words or phrases, or small sets of interdefinable words that appeal to important but small-scale frames. The Berkeley FrameNet Project (Fontenelle 2003, http://framenet.icsi.edu) is going beyond such piecemeal observations, and building a frame-based database containing hundreds of frames, many of which support quite large sets of words from the common vocabulary of English, accompanied by sentence annotations that serve both as the evidence for the analyses and as a collection of examples that can be made available for further research.

The method of inquiry is to find groups of words whose frame structures can be described together, by virtue of their sharing common schematic backgrounds and patterns of expressions that can combine with them to form larger phrases or sentences. In the typical case, words that share a frame can be used in paraphrases of each other. The general purposes of the project are both to provide reliable descriptions of the syntactic and semantic combinatorial properties of each word in the lexicon, and to assemble information about alternative ways of expressing concepts in the same conceptual domain.

13.2.1 General outline of the FrameNet process

The steps in the FrameNet lexical analysis process are as follows:

1. Characterizing the **frames**, i.e., the situation types for which the language has provided special expressive means.

> One of these is Revenge, the last phase of a scenario in which someone A had offended or injured someone B and after and because of that, someone C does something to punish A. The event designated in a Revenge predication is the punishing event.

2. Describing and naming the *frame elements* (*FEs*), i.e., the aspects and components of individual frames that are likely to be mentioned in the phrases and sentences that are instances of those frames.

> In the case of Revenge, the A of the previous paragraph is named the Offender, the B, the InjuredParty, and C, the Avenger. What the Offender did is referred to as the Offense; what C does is referred to as the Punishment. Phrases and clauses built around Revenge expressions are likely to mention some or all of these.

3. Selecting *lexical units* (*LUs*) that belong to the frames, i.e., words from all parts of speech that evoke and depend on the conceptual backgrounds associated with the individual frames.

> The long list of words that evoke the Revenge frame includes simple verbs like *avenge, retaliate, revenge*; phrasal verbs like *get back (at), get even (with), pay back*; nouns like *payback, reprisal, retaliation, retribution, revenge, vengeance*; adjectives like *vengeful, vindictive*; support constructions like *take revenge, wreak vengeance, exact retribution*, adverbial idioms like *quid pro quo, tit for tat*, and many others. Each LU is provided with a brief informal definition.

4. Creating *annotations* of sentences sampled from a very large corpus showing the ways in which individual lexical units in the frame allow frame-relevant information to be linguistically presented.

This is done with the help of annotation software that makes it possible for the annotator to associate FE labels with the phrases in a sentence that express those FEs. Example:

[$_{\text{AVENGER}}$ Hook] tries to [$_{\text{TARGET}}$ **avenge**] [$_{\text{INJUREDPARTY}}$ himself] [$_{\text{OFFENDER}}$ on Peter Pan] [$_{\textbf{PUNISHMENT}}$ by becoming a second and better father].

Grammatical functions (subject, object, etc.) and phrase types (NP, PP, etc.) are associated with the FE-tagged constituents by a mini-grammar in the software; these initial assignments are corrected by the annotator when necessary.

5. Automatically generating lexical entries, and the valence descriptions contained in them, that summarize observations derivable from them.

The above example would thus appear in the lexical entry with this information:

[$_{\text{AVENGER:SUBJECT:NP}}$ Hook] tries to [$_{\text{TARGET}}$ **avenge**] [$_{\text{INJUREDPARTY:OBJECT:NP}}$ himself] [$_{\text{OFFENDER:OBLIQUE:PP}}$ on Peter Pan] [$_{\textbf{PUNISHMENT:OBLIQUE:PP-GERUND}}$ by becoming a second and better father].

Tables including information from the full collection of annotations of the verb *avenge* show that in addition to the Offender, the direct object can also instantiate the Offense, such as *the insult, his brother's murder, their humiliating defeat*. In the large variety of lexical structures in the Revenge frame, the part of the Offender can be introduced with various prepositions: *take revenge ON. . ., get back AT. . ., get even WITH. . ., retaliate AGAINST. . .*

Thus, the kind of semantic analysis carried out in this work involves characterizing the situation types evoked by (classes of) lexical units, determining the kinds of participant roles (frame elements) needed to complete the details of instances of any such frame, and discovering and documenting the ways in which such elements are syntactically realized. The result of this work as of 2008 is a collection of frame descriptions (more than 800), an index of more than 11,000 LUs, a large collection of sentences annotated as illustrations of given LUs (more than 150,000), and lexical entries that include informal definitions (for human users) and tabular valence descriptions.

13.2.2 A sample frame: Compliance

In this section we exemplify a frame treated in FrameNet and the LUs that have been assigned to it. Words of different parts of speech are included all of which evoke in the language user's mind the kind of situation characterized by the frame description. The wordlist for this frame, as in many other cases, includes antonyms as well as synonyms, since, except for the polarity contrast, the basic elements of their annotations will be the same.

The verbal definition of a frame is formulated so as to include the names of the frame elements in the definitions in a way that displays their roles in the described situations. The definitions are for annotators and users to keep in mind. The FrameNet project does not currently provide formal definitions, simulations, paraphrases in terms of semantic primitives, alignment with any of the existing ontologies, or reduction to first order predicate logic or other formal representations.

The Compliance frame is defined as follows, where the capitalized bold-font words are the FEs:[3]

> The words in the Compliance frame evaluate the degree of faithfulness to some **Norm** (rule, standard, accepted practice) on the part of a **Protagonist** or an **Act** or a **StateOfAffairs** for which the **Protagonist** is responsible.

Examples of sentences that convey Compliance situations are as follows (frame-evoking expressions are in bold):

> The wiring in the computer room **violates** the current building code.
> You were not **in compliance** with the trespassing laws when you climbed over that fence.
> Do you faithfully **observe** the dietary laws?
> Did Jesus **break** the Sabbath?
> Russia is urged to **adhere** to the agreement.
> Being **compliant** with the HIPAA security regulations is not easy.
> Google is **in breach** of California privacy laws.

The entities mentioned in texts that contain predicators from the Compliance frame can be (a) an **Act**, like climbing over a fence, (b) a **StateOfAffairs**, like the state of the wiring in a computer room, (c) a **Protagonist**, a person or non-person agent, such as you, Jesus, Russia, or Google, and (d) a **Norm** like a code, agreement, regulation, or law. The primary argument in a Compliance predicate will be one of the first three; the Norm appears as a direct or oblique complement.

The LUs that belong in the Compliance frame include adjectives (*compliant (to), contrary (to), obedient (to), observant*); simple transitive verbs (*breach, break, circumvent, contravene, flout, follow, honor, keep, obey, observe, transgress, violate*); intransitive verbs with prepositional complements (*abide (by), adhere (to), comply (with), conform (to), submit (to)*); nouns morphologically related to verbs in the frame (*adherence (to), breach (of), compliance (with), conformity (to), contravention (of), non-compliance (with), obedience (to), observance (of), submission (to), transgression (of), violation (of)*).

[3] The frame definitions given here are not verbatim copies of the definitions given in the website.

13.2.3 The variety of frames

Here we sample a variety of frames to show the range of ideas that can be included in a single frame and the meaning distinctions that are left to the individual lexical units. It is possible to think of each LU as evoking its own frame, each of these inheriting those properties of the "mother" frame shared by them all.

Frame name	Definition	Example LUs	Comment
Adorning	something partially or wholly covers something else	*adorn blanket cloak coat cover deck decorate dot dress encircle encrust envelop festoon fill film garnish line pave stud wreathe*	These verbs differ in imagery, distribution of the covering substance, etc.
Attaching	someone causes something to be connected to (or disconnected from) another thing using some means	*adhere affix agglutinate anchor append attach . . . sew shackle solder staple stick . . . tack tape tether tie truss untie weld yoke*	These words differ with respect to the kinds of things that get connected to each other, the methods and instruments that are used
Biological area	a geographical area is defined by the biota in it; natural, not man-made	*bog bush copse desert fen forest glade grassland . . . taiga thicket tundra veld wold wood woodland*	
Change of phase	an undergoer changes phase (intransitive)	*condensation condense defrost evaporate evaporation freeze liquefy melt solidification solidify sublimation sublime thaw vaporize*	The LU meanings differ mainly as to the before and after states
Change position on a scale	something undergoes a change in the value of some magnitude	*advance balloon climb decline . . . skyrocket slide soar swell swing triple tumble*	These LUs differ with regard to directionality, speed, cause, path shape, etc.
Fluidic motion	the motion of a liquid	*bubble cascade churn course . . . splash spout spurt squirt stream trickle*	

13.2.4 Frame elements

The frame elements stand for those entities or properties which may or must be present in any instance of a given frame: in a sense, they stand for the things worth talking about once a frame has been entered into a conversation. It is not always obvious what they should be or how many there are.

Core and peripheral FEs

A distinction is made in FN classification between "core" FEs and "peripheral" FEs: there are clear cases and unclear cases in trying to draw this distinction. Clearly, FEs that are obligatorily expressed should belong to the core—but in some cases central concepts of a frame do not need to be expressed. In the case of verbs, FEs that get realized as nuclear syntactic relations such as subject and direct object, should also be part of the core—though there are grammatical constructions that sometimes make this unclear.[4] FEs that are expressed by phrases with lexically specific morphological marking should also belong to the core.

If a verbal LU evokes a frame that is a kind of Event, then the semantic roles that are always available for events should be peripheral: place and time specifications and various other circumstantial notions. If an LU evokes a frame that is a kind of IntentionalAct, then phrases that indicate intentions, purpose, the actor's attitude, and the like, can be described as peripheral.

The trio "time, place, and manner" usually covers what grammarians mean by peripheral adjuncts rather than core arguments, but each of these semantic types can have core status in some lexical contexts. For example, a locative is required with the verb *reside*, as in *they reside in the eastern part of town*, and a manner complement is required with the verb *phrase* in the Encoding frame, as in *he phrased his reply in an exceedingly rude manner*. Utterances like **they reside* and **he phrased his reply* are not complete.

Some prepositions are selected by the LU. They indicate that the FE expressed by the PP is core: *interested in X, familiar with X, hostile to X; adhere to X, depend on X, accuse NP of X; relation to X, pride in X, hostility to X*; and so on. Other PPs express peripheral notions and do not vary in ways selected by the lexical head: *in the kitchen, on Thursday*, and *in a hostile manner* have constant forms wherever their meaning is needed. The same PP can be one or the other depending on its context: *on the bus* can refer to the location of some person (*I met my girlfriend on the bus*), because in this case the preposition is selected by the noun, or it can be a core FE as in *we can depend on the bus*.

For situations involving visual and auditory perception it may be necessary to recognize a distinction between a locating expression that indicates the location of the entire event (*I saw Harry in New York last week*) and one that indicates the

[4] For a sentence like *I'll bake you a cake*, the apparent direct object is not a straightforward entity in the baking event, but there is a benefactive construction behind this sentence for which the "benefited" element is necessary; similarly, *I ate my plate clean* is the product of a resultative construction, and the plate is not to be interpreted as a FE of the ordinary verb *eat*, but is an FE of the complex construction *eat clean*.

location of the perceived entity alone (*I saw the prowler on the neighbor's roof*). Expressions in which a locating expression gives the location of the percept may co-occur with expressions which specifically locate the perceiver (*from my kitchen window I could see the prowler on the neighbor's roof*). Such distinctions are not recognized for ordinary transitive verbs like *eat*, *fold*, *squeeze*, etc.

Frame element naming

Frame element names are defined relative to individual frames; in a sense this means that the names need to be chosen so that someone who understands the frame will be able to see what FEs the names refer to, and so that annotators will easily keep in mind what they are looking for.

This practice distinguishes the FrameNet procedure from two extremes: one that limits semantic role names to a small fixed set intended to cover all cases,[5] requiring that the same names have the same interpretations wherever they are used; and the other extreme chooses names that are specific to individual lexical items (the two main arguments of *see*, for example, could be called the Seeer and the Seeee!). The theoretical importance of FE names is only that they be distinguished from each other in describing an individual frame.

The practical value of frame-specific names over generalized semantic role names is that annotators can have case-specific names in mind when they are tagging the phrases that represent particular FEs, rather than trying to fit very abstract concepts to local cases. The advantage of using frame-specific names over LU-specific names is that comparable names can be re-used when annotating sentences from the same frame, and a mechanism for computing potential paraphrasability can be developed by comparing FE names in sentences sharing the same frame.

Extrathematic FEs

Since FrameNet annotation is aimed at accounting for all of the constituents that occur in phrases built around a frame-evoking target, there is one more type of semantic element that shows up in such sentences. Some sentence adjuncts are said to be *extrathematic*, in the sense that they introduce information that is not a necessary part of the description of the central frame. In many cases such phrases introduce a new frame and in a sense attribute that frame to the rest of the sentence. For example, in a sentence like *He set fire to his ex-wife's car in retaliation*, we know that the concept of Revenge, introduced in the phrase *in retaliation*, is not a part of the action of igniting something; in this case the main clause is itself interpreted as the means by which revenge is achieved.

[5] The search for this "minimal set" has continued for decades, e.g., Fillmore (1968: 24–5), Frawley (1992: 197–249), Jackendoff (2002: 143).

13.2.5 Syntactic contexts for FE realization

The lexicographic work of locating the FEs that fill out the frame associated with an LU needs to be distinguished from other ways of finding information in a text. One paradigm for analyzing text is that of information extraction, in which any kind of information available in a text can be drawn upon for filling out pre-existing templates. In the "Dream Breakfast" text, for example, a text interpretation algorithm might seek to fill in a template about the breakfast experience, determining that the Eater is the woman referred to throughout with feminine pronouns, that the Food includes wheaties and milk, that the Beverage is coffee, and that the Location is a table in the garden. This kind of process combines information across sentences and requires anaphora resolution, analysis of the sequences of events, etc.

The frame semantic approach differs from information extraction in seeking to develop descriptions of the LUs based on their combinatorial properties. Thus, in identifying FEs in sentences, it is necessary to notice (1) the particular LU that evokes the frame, and (2) the positions in the sentence that are "in grammatical construction" with that LU that provide frame-elaborating information. In our case we are concerned with the combinatory affordances of individual lexical units in the language, i.e., what grammatical positions they provide, what must or what need not be expressed, and so on.

This is worth exploring, because the attempt to detect frame-relevant information in grammatically relevant positions has led to moving beyond ordinary lexical complementation. The familiar notion of subcategorization frame gives an obvious place for satisfying FE requirements for a lexical item, namely after the LU within the phrasal projection of that LU, i.e., inside the VP for a verb, inside the NP for a noun, inside the AP for an adjective, and so on. Subjects of finite sentences are also FEs of the predicates heading the sentence, as the bracketed elements in these sentences. FEs detected for FrameNet purposes require going beyond these cases.

Some adjectives are treated as evoking frames in FrameNet, but others are treated mainly as satisfying FEs for frames evoked by the nouns they modify. *Descriptive adjectives*, when used attributively, realize one of their FEs in the noun that they modify: suspiciousness is attributed to the modified constituent in *suspicious [behavior]* and in *[something] suspicious*. Descriptive modifiers of nominals assign FE-status to their heads, where the head is not treated as a frame-bearing element in its own right. By contrast, *relational adjectives* generally serve to identify an FE of the frame evoked by the nouns they "modify": these are the adjectives typically defined with the phrase "of or pertaining to ...". A *policy* has to cover some domain: an *economic policy* is a policy whose domain is "the economy"; a *law* has to cover some area of compliance: *criminal law* is the kind of law that deals with crimes. Some adjectives have both descriptive and relational functions, as seen in the distinction between *educational policy* and *an educational experience*: the former is a policy about educational practice, the latter is an experience from which something can be learned.

In N + N compounds where the second noun is a frame-evoking LU, the first often expresses one of that noun's FEs. Thus, *homeland security* has to do with keeping the homeland secure; *health risk* is a risk to your health; *fire risk* is the risk of fire. FEs of noun LUs that designate events and relationships may appear as the possessive determiner of the noun. Thus, *my decision* is something that I decided; *your dismissal* is an event in which someone dismissed you; *his similarity to you* can refer to the way in which he is similar to you.

In the examples above, the fillers of the FEs have all been in direct structural relation to the relevant LUs. However, syntactic theory provides other ways to locate fillers of the FEs of a frame-bearing LU by means of argument sharing. This includes arguments shared with higher "embedding" verbs, nouns, or adjectives. For the arguments of non-finite verbs, this will include shared arguments with "higher" embedding verbs (*[Pat] tried to* ***kill*** *himself, [the deer] seems to* ***understand*** *us, [the letters] keep* ***coming***, *[this book] is worth* ***reading***), or adjectives (*[the enemy] is likely to* ***surrender***, *[I] have been eager to* ***meet*** *you*), or nouns (*[Pat's] decision to* ***resign***). Perhaps less familiar are *support constructions*, which offer FE-information about a frame-bearing event-noun or relational noun in the syntactic arguments of a governing verb. Those support structures that exist mainly for giving verbal expression to a noun's meaning, without adding much semantic information of their own, are the so-called *light verbs*: *pay attention, give advice, take a bath, have a disagreement*, etc., where the subject of the verb is the one who attends, advises, bathes, disagrees, etc. In other cases, the subject of a support verb provides information about a secondary participant in an event: *undergo an operation*, *get advice*, and so on. Going beyond light verbs, there are other argument-sharing structures that add registral information (*to lodge a complaint, to issue a decree, to submit an application*); there are some that share arguments with the noun, but describe secondary events related to the event named by the noun (*to break a promise, to pass a test, to keep an appointment*).

13.2.6 Null instantiation

Despite the large number of structural positions in which FEs can appear, sometimes core FEs of an LU are simply not to be found anywhere in the sentence built around the LU. There appear to be three main explanations for unrealized FEs, one grammatical and two lexical. The grammatical explanations have to do with structures that require or permit the omission of some argument position. This is referred to as *constructional null instantiation*. Examples are imperative sentences (*Please leave the room*), where the omission of the subject is licensed by the grammar, and passive sentences (*We've been fooled*) where omission of the agent phrase is permitted.

The other two are called *indefinite null instantiation* (INI) and *definite null instantiation* (DNI), lexically licensed zero anaphora. Instead of declaring that the verbs *eat* and *win* can be either transitive or intransitive, FN regards them as always transitive, but records the fact that the Food argument of the verb *eat* can be omitted under INI (*I've already eaten*) because the situation implies that what was eaten does not matter. On the other hand, in *we won*, the Contest FE has been omitted under DNI, implying that all parties to the conversation know what it is. Those were both examples of missing direct objects; the same contrast can be found with PPs: if someone says *That depends* (INI), you won't know what the Contingency is that the speaker has in mind unless you ask; but if someone says, *When did they arrive?*, it is assumed that you already know what Destination they had in mind. Both DNI and INI have implications for text coherence, as will be shown in Section 13.4.4.

13.3 Frame-to-frame and FE-to-FE relations

FrameNet implements a large proportion of the concepts of Frame Semantics, with many of the concepts directly mapped to the database structure. For example, there is a table for frames, and another for frame elements, with a reference from the FEs to the frame they are part of; the table of lexical units is likewise linked to the frames and to the lemmas, representing the meaning and the form sides, respectively. As the number of frames has grown, it has become obvious that they are not simply a collection of separate entities, but there are networks or hierarchies of frames, that some are instances of others, some are components of others, etc., and so an important part of the FrameNet project has been to work out this system of relations.[6]

13.3.1 Types of frame-to-frame relations

The FN frames are now linked by a system of several hundred frame-to-frame relations, which allows assertions about semantic types to be made at the appropriate

[6] Much of the effort in working out the details of these relations has come from FrameNet staffer Michael Ellsworth.

level of generality. There are eight relations, one of which is not relevant here;[7] the others fall into three groups, Generalization, Event structure, and Systematic.

(1) **Generalization relations**:

- **Inheritance**: All FEs of the parent frame are bound to FEs of the child frame, but the child FEs need not have the same name. The child can have more FEs, and the child's semantics is a subtype of the parent's semantics. For example, the Revenge frame inherits from the Rewards and Punishment frame, since it also involves a person inflicting a punishment on another. It differs explicitly from Rewards and Punishments in being outside of institutional or judicial control.
- **Perspective on**: Different lexical items (e.g., *buy*, *sell*) evoke frames with different perspectives on an abstract event (Commercial transaction), a kind of figure:ground relation (Gawron 1988: 151ff.). Specifically, buying takes the perspective of one participant in the goods-transfer, and selling takes the perspective of the other. In FN, they are linked to the abstract event via Perspective on relations.
- **Using**: The child frame depends upon background knowledge provided by the parent frame; at least some of the core FEs of the parent are bound to child FEs, but not all of them.

(2) **Event structure relations**:

- **Subframe**: These are sub-events of a complex event, often with temporal ordering, e.g., in FN the Giving frame is linked to two sister frames, called, for lack of better names, Pre-Giving and Post-giving, which provide information about who has what when (and inherit, indirectly from the Possession frame). These three frames together constitute the Giving scenario.
- **Precedes**: This relation specifies temporal ordering, e.g., Giving Precedes Post-giving.

(3) **Systematic relations**:

- **Causative of**: The parent frame represents the causative corresponding to the child frame, e.g., Cause change of position on a scale (LUs *raise, lower*) is the Causative of Change position on a scale (*rise, fall*).
- **Inchoative of**: The parent frame represents the inchoative, and the child represents the stative. Change position on a scale (*rise, fall*) is the Inchoative of Position on a scale (*high, low*).

Note that *all* of these frame-to-frame relations have accompanying frame element-to-frame element relations (including some not detailed above).[8] Also, there is a

[7] The "See also" relation, which redirects the reader to another frame, much as dictionaries often contain cross-references to other words.

[8] Because the frame and FE relations form a very complex graph, a tool called "frame grapher" has been provided on the FN website to make it possible to browse the graph.

system of semantic types which apply to the FEs, and are also inherited, so that most of the FEs have both semantic types and a relation to higher-level FEs; those near the top of the hierarchy are similar to traditional thematic roles (case roles), such as Agent, Theme, Source, Goal, etc. Where there are FEs that do *not* map to these higher-level roles, it is deliberate—there are simply no general semantic roles that apply to the specific situation which the frame embodies. Every instance of a frame or a frame element is also an instance of the more general frames or frame elements from which it inherits.

13.3.2 Commerce: A detailed example of frame relations

The full network of frame-frame and FE-FE relations is now capable of representing the semantics of ordinary events in considerable detail. We will exemplify this with buying and selling events, which have received much attention in the lexical semantic literature (Fillmore 1977*b*; 1985; Fillmore and Atkins 1992; Gawron 1988; to appear). For the most part, FN adopts a simple three-state model of event structure, with a pre-state, a central change, and a post-state, although much more complex events can be modeled, as we shall show. Typically, only the central, changing part of the event is profiled, and, hence lexicalized; for example, a getting event involves a pre-state in which the protagonist doesn't have something, the central change, lexicalized with the verb *get*, and a post-state in which the protagonist has something. In FN, these stages are all implemented as related frames, as shown in Figure 13.1, in which ovals represent non-lexical frames and rectangles, lexical frames.

We also describe all three stages as subevents of a more complex event, called a Getting_scenario (also implemented as a non-lexical frame), related to them through the Subframe (Sub) relation, as in Figure 13.2.

Buying is a type of getting, specifically, getting goods and services in exchange for money. It represents one part of a commercial transaction, which profiles the buyer's acquisition of the goods; an alternative perspective on the same event is that of the seller, which profiles the seller's giving of the goods. We therefore say that the frame Commerce_buy Inherits (Inh) from Getting and that Commerce_sell

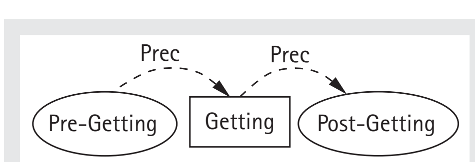


Figure 13.1. Generalized three-state event model

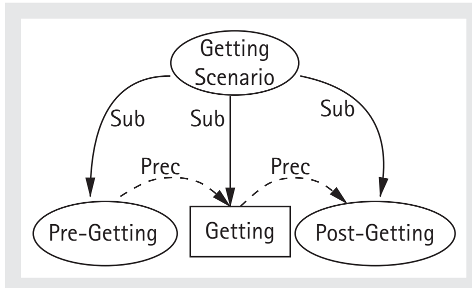


Figure 13.2. Three-state event model as subframes of an event scenario frame

Inherits from Giving. Both have to do with the transfer of goods, but each has a different Perspective_on (Perspect) relation to a more general Commerce_goods_transfer frame, shown in Figure 13.3.

However, the commercial transaction has two subframes (subevents), the transfer of goods and the transfer of money. In this case, unlike the simple event frame, there is no fixed temporal order; the two transfers can be simultaneous, or either can precede the other, so we simply do not create any precedes relations between the two subframes. Again, although the Commerce_money_transfer frame itself is non-lexical, there are two lexical frames, Commerce_pay and Commerce_collect which represent different perspectives on the transfer of money, that of the buyer and that of the seller, and paying is a form of giving and collecting, a type of getting, so there is a partial symmetry between the two types of transfer, as shown in Figure 13.4.

In each case, the inheritance from the Giving and Getting frames helps explain the valence patterns; because buying and collecting are getting, we buy *from* the

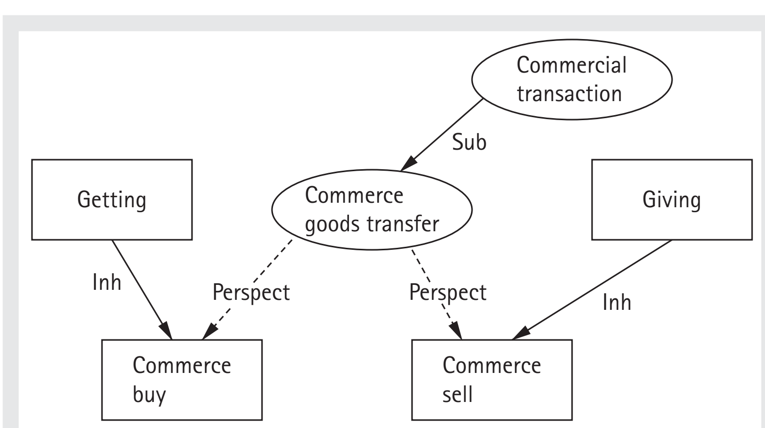


Figure 13.3. Frame relations around Commerce_goods_transfer

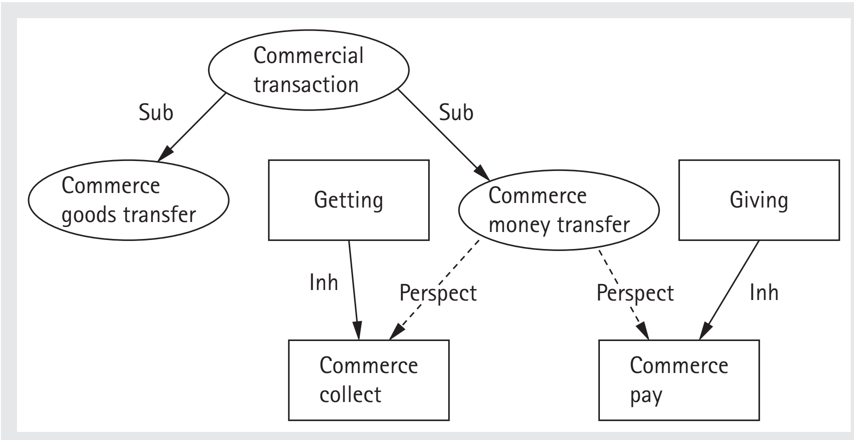


Figure 13.4. Frame relations around Commerce_money_transfer

seller and collect from the buyer, and because selling and paying are giving, we have the dative alternation: *Chuck sold Jerry a car/Chuck sold a car to Jerry* and *Chuck paid a dollar to Jerry/Chuck paid Jerry a dollar*.

13.4 A Frames perspective on familiar issues in linguistic semantics

Since information in FrameNet is organized by frames, there is no direct representation of LU-to-LU relations in the database—relations such as hyponymy, synonymy, antonymy, or polysemy.

13.4.1 Polysemy

A frames approach to lexical semantics, by which an LU is a pairing of a lexeme with a sense (and that usually means the pairing of an LU with the frame in which its meaning is defined), is more or less required to take a "splitting" rather than a "lumping" approach to polysemy. This is because an LU is located in a system of relations (a) with its combinatorial properties and (b) with other words in its frame. This means, for example, that, whatever features are shared by the word *short* when it is paired with *long* (and used in measuring unoriented extents) versus when it is paired with *tall* (and used in measuring vertically oriented extents), two LUs are involved. This also means that the two uses of temporal unit words like *year*, *month*, *week*, *day*, and *hour*, are going to be separated according to their *calendric* uses (where they designate calendar or clock cycles with clear beginnings and endings) and the *metric* uses (where they designate temporal spans measured by the length of such a cycle. Simple phrasing differences select the one sense or the other: if it is Thursday and you are told that your project must be completed soon, the urgency is different depending on whether the completion must be *within a week* (you have seven days) or *within the week* (you'd better hurry). The expressions with *the* select the calendric sense of each of these words.

A study of the range of senses that can be attributed to a single lexeme then has to be done by finding lexemes of the same form in more than one frame. The Compliance frame provides two useful examples of the importance of the LU concept. In particular, the lexeme *adhere* appears both in Compliance and in what is called the Attachment frame, having to do with things connecting to each other (*the bandage adheres to the wound*). While it could be argued that the *adhere* of Compliance is a *motivated* sense built on the Attachment sense ("one attaches one's behavior to some principle"), evidence that the two LUs *adhere* belong in different frames is found in their morphological relation to the corresponding nominalizations. By separating them we can say that the nominalization for the Compliance sense of *adhere* is *adherence*, while the corresponding noun for the Attachment sense is *adhesion*. A second example in the Compliance frame can be seen with the verb *observe*: this word belongs to a Perception Passive frame as well as the Compliance frame: its nominal derivative in Compliance is *observance*; the noun in the Perception Passive frame is *observation*. One of the advantages of separating LUs with the same form is that facts about meaning, valence, and other distributional properties, as well as correspondences between word forms sharing a morphological base, can be stated at the level of the LU and not the lexeme.

It occasionally occurs that the same lexeme supports more than one LU in the same frame, especially in the case of two kinds of nominalizations. In the Replacement frame, which concerns a situation in which something New takes the place of something Old (or, in the associated causative frame, some Agent causes something New to take the place of something Old), the nominalization *replacement* has two senses, one referring to the *process* by which a Replacement event takes place, and one referring to the New item itself, i.e., the *product* of a Replacement event. (*The replacement of Bunt with Schmidt came as quite a surprise. Schmidt is going to be Bunt's replacement in the next match.*) In the case of a different verb in the Replacement frame, *substitute*, there are distinct nominalizations covering the analogous senses: *substitution* for the process, and *substitute* for the product. (*The substitution of Schmidt for Bunt came as quite a surprise. Schmidt is going to be Bunt's substitute in the next match.*) In such cases, of course, the LUs are not defined solely as a pairing of a lexeme with a frame.

13.4.2 Antonymy and negation

FN does not currently have a treatment of the many uses of negation, but since antonyms are generally included in the same frame, inferences based on kinds of antonymy should be possible. For all frames in FN that project scalar interpretations, those members of the frame that have a "negative" orientation are marked with Negation as a semantic type. At present nothing has been done to take into account the variety of notions of opposition in lexical semantics, but users of the database for purposes of paraphrase generation need to include such information in their computations. In the Compliance frame, for example, compliance and non-compliance are *contradictory*, so some act which is *compliant with* a law is *not in violation of* the law.

Where antonyms are *contrary*, weaker inferences are possible: if something is increasing, it is not decreasing, if it is decreasing it is not increasing, but the opposite inferences are not possible.

13.4.3 Synonymy and paraphrase

Synonymy proper, if such a relation exists, would theoretically be represented by words of the same part of speech in the same frame, with the same valences and with the same definitions. Since FN has only informal definitions, it offers no precise way of identifying synonyms. By being frame-based, however, it does allow for the recognition of *paraphrase* relations. As a tool for use in automatic question-answering and information extraction, the possibility of generating paraphrases for incoming questions should be enhanced by the ability to generate sentences with the same frame elements but different lexical-syntactic forms.

Simple cases: Many instances of paraphrase involve part-of-speech alternating locutions that include the same FEs. Thus, within Compliance, the following would be considered paraphrases: *This conforms to the regulation/is in conformity with the regulation/is compliant with the regulation*. Here different words and phrases from the same frame associate the same relationship between a State-OfAffairs and a Norm.

Negation cases: In cases of contradictory relations between antonyms in a frame, positively vs. negatively formulated paraphrases are possible. *This conforms to the regulation/does not violate the regulation. This is in violation of the regulation/is not in compliance with the regulation.*

Perspectivally related cases: Pairs of expressions in a complex frame-tree may present different perspectives or profilings on a single type of event or relationship, as we have seen in the case of commercial transaction verbs. *A sold B to C* relates the same event as *C bought B from A*. Similarly, *A employs B* describes the same situation as *B works for A* though with a different perspective.

Inheritance related cases: Particularly important in this regard are paraphrases involving frames that have an *inheritance* relation between them. If one frame inherits something from a more abstract frame but adds something new, then paraphrases can be created using the two frames if an extrathematic element introduces the special element into the higher frame. The discussion above of frame-to-frame relations showed that buying is a kind of getting, in particular, getting in exchange for a money payment. Extrathematic reference to the payment in a getting expression could then compositionally be equivalent to a similar expression with buy: *I got it for $12/I bought it for $12*. The *for*-phrase is extrathematic in the former sentence, thematic in the latter.

Systematic relations cases: The relations of Causative and Inchoative make it possible to recognize paraphrases between synthetic and analytic expressions of these concepts, as in *We made it higher* vs. *we raised it*.

13.4.4 Coherence and anaphora

The "texture" of a linguistic text (Halliday and Hasan 1976) is the system of meaning links from one part to the next, and from parts to a whole, and one of the major tools in showing the need for such coherence links is the system of implicit arguments identified in FrameNet as definite and indefinite null instantiation. Usual anaphora resolution schemes operate on what are called "mentions"—words or phrases in a text that mention the same entity, and the resolution problem is to show chains of mentions that all refer to the same thing. Recognizing lexically licensed null instantiation offers new challenges to anaphora resolution. On encountering a sentence like *My explanation is quite similar*, one would know that the preceding discourse had introduced a problem or mystery (*explanation* is missing one of its arguments), and that somebody had offered an explanation of that mystery (*similar* is missing one of its arguments), and that the speaker is now likely to offer his own explanation. The *of*-phrase that could be expected with the noun *explanation* (*explanation of the mystery*) is omissible only when it is known in the context; the *to*-phrase that could be expected to accompany the adjective *similar* (e.g., *similar to your explanation*) is omissible under similar conditions: all participants are aware of the previously topical explanation. One FE of *similar* that is also missing is of the INI sort, namely the parameter in respect to which of the two explanations are similar. It is likely that the next utterance of the speaker of this sentence is going to be a description of that similarity.

13.5 Beyond English: FrameNets in other languages

Since frames are defined on semantic grounds, we expect most of them to be comparable across languages; e.g., the concept of a commercial transaction will be much the same in any culture, although details may vary. Other frames, such as the stages of the criminal process (indictment, arraignment, bail-setting, etc.), are more culture-specific.[9] As of October 2008, several projects using a Frame Semantic approach for annotating languages other than English had already begun to bear fruit and several others were just starting. (We will abbreviate the English FrameNet in Berkeley as "BFN".)

The SALSA Project (Burchardt et al. 2006, http://www.coli.uni-saarland.de/projects/salsa/) is annotating German newspaper texts using BFN frames and FEs. For most German words, they find an appropriate BFN frame and FE labels; if nothing suitable is found, they create predicate-specific "proto-frames". SALSA I began in 2003, and released a first version of the data, of roughly 20,000 annotation sets. SALSA II is investigating how the proto-frames and LUs created in SALSA I might be integrated into a German FrameNet and/or the English FrameNet.

The Spanish FrameNet Project (Subirats and Sato 2004, Ellsworth et al. 2006; Subirats 2007, http://gemini.uab.es:9080/SFNsite) has been developed at the Autonomous University of Barcelona by Carlos Subirats since 2002, using the BFN annotation and report software. They plan a formal data release in 2008, comprising 1,000 LUs over a wide range of semantic domains.

Japanese FrameNet (Ohara et al. 2004; Fujii 2005, http://jfn.st.hc.keio.ac.jp/), led by Kyoko Ohara of Keio University with colleagues at Keio University, Tokyo University, and Senshu University, has been under development since 2000, first building their own corpus and extraction tools for Japanese and then modifying the BFN software to handle Japanese. JFN worked first on verbs of communication, then on motion and perception nouns and verbs. Hiroaki Sato, of Senshu University, Kawasaki, Japan, has built a system called FrameSQL (http://sato.fm.senshu-u.ac.jp/fn23/notes/index2.html) to search the FrameNet data according to a variety of criteria. As FrameNets in other languages have grown, he has also devised ways of aligning LUs across languages using a bilingual dictionary (Sato 2008); where the same frame is found in both languages, one can view annotated examples of the same frame and FEs for corresponding LUs in two languages.

Chinese FrameNet has been underway at Shanxi University in Taiyuan since 2004, using their own corpus and corpus search and annotation software (You and

[9] For further on frame semantics across languages, see Ellsworth *et al.* 2006, Boas 2005, and Lönneker-Rodman 2007.

Liu 2005, You et al. 2007). The CFN database now contains more than 2000 LUs in more than 300 frames, with more than 20,000 manually annotated sentences; as elsewhere, most of the LUs fit in BFN frames.

Alessandro Lenci, of the Department of Linguistics at the University of Pisa, and Martina Johnson began an Italian FN in 2008, starting with communication verbs, and using a copy of the BFN database and software. Birte Lönnecke-Rodman set up a FN database for Slovenian in 2007, inserting a Slovenian lexicon, and has been able to annotate and produce reports from the annotation. Efforts have also started on establishing a Brazilian Portuguese FrameNet, led by Margarida Salomão, of Universidade Federal de Juiz de Fora.

13.6 Beyond the lexicon

The Berkeley FrameNet Project has been devoted to the frame structures of lexical items and how information about individual frame instances is provided in sentences built around the lexical items studied. This has not included work on negation and quantification, central to much work in formal semantics, though there is no obvious reason why a frames approach cannot be extended into such areas: a corpus-based approach that maximized computational techniques has not appeared sufficiently subtle for the kinds of phenomena noticed in such studies.

The current database deals only with lexical units, while recognizing of course that many lexical units are themselves made up of phrases or other kinds of (possibly discontinuous) word groupings. There are linguistic properties smaller or more abstract than lexical items that contribute greatly to the semantic organization of sentences—number, tense, aspect, modality, and the closed-class categories and morphemes—as is well documented in the work of Langacker (especially Langacker 1986) and Talmy (especially Talmy 2000).

There are also grammatical constructions which evoke semantic, pragmatic, or interactional frames on their own. The various ways of forming commands, questions, wishes, curses, threats, etc. involve understandings about how the participants in the ongoing conversation are interacting with each other. The same kind of frame analysis that can treat *She ordered him to leave the room* should be adaptable to *I order you to leave the room* and to *Leave the room!* In other words, the participants in a linguistic frame can easily be participants in the communication event itself.

Beyond this, a large number of minor grammatical constructions bear very special interpretation requirements. For a simple example we may take the use of the quantifier *no* (or its variant NEG[10] +*any*) accompanying a comparative phrase in predicating position. There's a striking difference between a simple negation (with *not*) of a comparative phrase, as in the neutral description *it wasn't bigger than a golf ball*, on the one hand, and the marked expression *it was no bigger than a golf ball*, indicating that the speaker regards a golf ball as quite small (and is therefore expressing the idea that the object in question is also quite small). In this phrasing only the unmarked adjective is selected, in the case of scalar opposites.

Less mysterious constructions are the double-NP means of expressing rates, where the first NP identifies a quantity of one kind of unit and the second identifies a different kind of unit or measure: examples are *two milligrams a day, twenty miles an hour, a hundred miles a gallon, 300 times per second, twice every three days*, and the like. Nothing else in the grammar of English predicts the manner of interpretation of these expressions. They may be iterated: *twenty milligrams twice a day, $300 a person per day*. Compositional semantic principles operate on the structure of the phrases and define such notions as Frequency, Speed, Dosage, Cost-per-Unit, Growth Rate, and many others, usually serving as FEs of some frame in their environment.

In addition to phrasal patterns with special interpretations, there are a great many cases in which individual words that "start out" in one frame are used in a context which places them in another frame, given regular or semi-regular interpretation principles that relate the two framings. This includes nouns that acquire different status respecting the count/noncount distinction (*we had beaver for lunch* [animal name as meat from such an animal], *we enjoy the wines of Rioja* [mass noun as countable variety of that mass]), and several others. Richly described in current linguistic literature are variations in verbal argument structure, by which a verb indicating one kind of activity is used with the valence typically associated with another kind of activity usually communicating that the former is a component of the resulting event type, as in *eat your plate clean* or *she sneezed the napkin off the table* (Boas 2003; Goldberg 1995; 2006).

Future FrameNet activities will be moving into the semantics of grammar, both general and abstract (negation, tense, aspect) and phraseological (constructions and syntactic idioms), making it possible in principle to test methods of integrating lexical meanings and grammatical meanings into a complete account of the language-based interpretations of texts.

[10] The NEG can be the word *not* or other negative polarity determining contexts. Compare: *you're no better than he is, you're not any better than he is, I doubt you're any better than he is.*

CHAPTER 14

FRAMEWORK-FREE GRAMMATICAL THEORY

MARTIN HASPELMATH

FRAMEWORK-free grammatical description/analysis and explanation is argued here to be superior to framework-bound analysis because all languages have different catgeories, and languages should be described in their own terms. Frameworks represent aprioristic assumptions that are likely to lead to a distorted description of a language. I argue against restrictive theoretical frameworks of the generative type, against frameworks of functional approaches such as Functional Grammar and Role and Reference Grammar, and against Basic Linguistic Theory.

14.1 Why framework-free?

While some readers may find this surprising, in this chapter I claim that there are many linguists who carry out theoretical research on grammar but do not work

I am grateful to Frederick Newmeyer, Edith Moravcsik, Bernd Heine, and Heiko Narrog for useful comments on this chapter.

within a theoretical framework, and I show how this is done. As far as I know, this theoretical stance has not been articulated in detail before, at least not in contrast with the typical 20th-century frameworks, some of which are represented in this book. There is a widespread view that it is in principle impossible to do framework-free grammatical research, and that those who do not adhere to a particular framework in fact work in an "eclectic" framework or in the framework of "Basic Linguistic Theory". I will argue here that this is not the case. Framework-free grammatical theory is not only possible and widely practiced but is, I believe, the best approach to the scientific study of language structure, though of course the space limits of this chapter do not allow me to make a full case for this assertion.

Most linguists seem to agree that we should approach any language without prejudice and describe it in its own terms, non-aprioristically, overcoming possible biases from our native language, from the model of a prestige language (such as Latin or English), or from an influential research tradition (such as that of Donatus's Latin grammar, or Chomsky's generative grammar). I argue that this is absolutely essential if we want to come even close to doing justice to our research object, and that, moreover, any grammatical framework is precisely such a "prejudice" that we want to avoid. Frameworks set up expectations about what phenomena languages should, can, and cannot have, and once a framework has been adopted, it is hard to free oneself from the perspective and the constraints imposed by it. What we need instead is the researcher's ability to discover completely new, unexpected phenomena, to detect previously unsuspected connections between phenomena, and to be guided solely by the data and one's own thinking.

One might object that, while this is a noble goal, it is in fact impossible, and that it is better to adopt some off-the-shelf framework and work within it, even if one is aware of some of its limitations. Against this, I argue that framework-free theorizing is possible and that it is practiced more widely than many linguists think. But before we can get to some concrete examples, a few key concepts need to be discussed in the next section (14.2). (Readers with time constraints may skip section 14.2 and read it only at a later stage, to allow them a deeper understanding of the main points and the terminology adopted here.) In section 14.3, I argue for and exemplify framework-free grammatical analysis, and in sections 14.4–6 I discuss the problems associated with three kinds of frameworks, restrictive frameworks (section 14.4), functional frameworks (section 14.5), and Basic Linguistic Theory (section 14.6). In addition to grammatical analysis, grammatical theory also has comparative and explanatory tasks, and the framework-free approach to these is presented in section 14.7.

14.2 Some fundamental concepts

14.2.1 Grammar

My topic here is grammatical theory, i.e., theory of morphosyntax. Very similar issues arise in phonology, but I will not discuss phonological frameworks and framework-free phonological theory here (but see Mielke 2008 for a recent account of phonology that is very similar in spirit).

14.2.2 Framework

A framework (also called *descriptive framework* or *theoretical framework*) is a sophisticated and complex metalanguage for linguistic description that is intended to work for any language. As Dryer (2006*a*: 29) notes, it is often possible to "translate" a particular analysis from one framework into another framework (e.g., from Relational Grammar into Government-Binding Theory), as is expected if frameworks are metalanguages. Such translations are often not completely equivalent, that is, the two analyses are more than notational variants of each other. But since descriptive frameworks tend to be complex and difficult to master, and few linguists specialize in translating between frameworks, it is often difficult to see which aspects of an analysis are specific to a particular framework and do not translate readily.

Descriptive frameworks are often called *theoretical frameworks* or simply *theories*, but this practice is not followed here because the term *theory* has multiple senses and is best reserved for another sense, as we will see in the next subsection.

14.2.3 Theory

I distinguish four senses of the term *theory* here, all of which are common in current linguistics. I find it most useful to limit the application of this term to senses 3 and 4. The term *theory* in the title of this chapter is intended in sense 4.

Sense 1: As we saw in the preceding subsection (14.2.2), *theory* is often used in the sense "descriptive framework" for a sophisticated metalanguage for describing languages.[1] Some of these frameworks have *theory* in their name (e.g., Government-Binding Theory, Optimality Theory, Basic Linguistic Theory). Framework-free

[1] Cf. Dryer (2006*a*: 28–9): "The notion of theory widely assumed in formal linguistics is essentially equivalent to that of a metalanguage for describing languages. Providing an analysis of a particular set of data within a formal theory involves providing a description of that data within the metalanguage that constitutes that theory."

descriptions are sometimes seen as "atheoretical", and this is correct if *theory* is used in sense 1.

Sense 2: A theory is sometimes understood as an abstract model or description of a complex empirical domain. Thus, one can say that a description of English is a theory of the competence of an English speaker.

Sense 3: A theory can be a set of coherent hypotheses or claims about a particular phenomenon, e.g., a theory of what caused dinosaurs to die out, or a particular theory of restrictions on *wh*-movement.

Sense 4: Finally, the term *theory* can be used in a loose sense, referring to theoretical (i.e., non-applied) scientific work, or "theorizing". It is in this sense that *usage-based theory* and *valency theory* should be taken in this handbook, and it is in this sense that *theory* is used in the title of this chapter.

Thus, in this chapter I discuss theorizing about morphosyntactic phenomena that makes no use of descriptive frameworks.

14.2.4 Description

By *description* I mean the characterization of grammatical regularities of particular languages. Grammatical descriptions must make use of abstract general entities such as rules, schemas, and constraints, because all languages allow an indefinitely large number of sentences and it is therefore not possible to describe a language by listing all its sentences.

It is often said that linguists should strive not only to describe the rules in such a way that speaker behavior can be predicted accurately ("phenomenological description" in Haspelmath's 2004 terms) but they should also strive to describe languages in such a way that the description reflects the speakers' internal generalizations correctly ("cognitive description", or "descriptive adequacy" in Chomsky's terms). However, it is far from clear that the latter is an attainable goal because often different generalizations are compatible with the facts, and we have no way of knowing which generalization is adopted by the speakers (note that it could be that different speakers have different generalizations). Thus, linguists must by and large be content with descriptions that accurately predict the behavior of speakers in natural corpora and experimental contexts.

14.2.5 Analysis

I use the term *analysis* synonymously with *description*. In linguists' current usage, *analysis* generally seems to imply a higher level of generalization, but this is a matter of degree. All linguistic description must involve generalizations (rules, schemas, constraints), and there is no distinction in principle between shallower and deeper generalizations. (Another usage of the term *analysis* is in the sense "description within a particular framework". Many papers in the generative tradition first provide a fairly framework-free description of the relevant phenomena ("the data") and then go on to provide a second, framework-bound description ("the analysis"). Since this chapter argues against framework-bound descriptions, this second sense of the term *analysis* is not of interest here.)

14.3 Framework-free grammatical analysis

14.3.1 Advantages

Most linguists agree that in describing or analyzing an unfamiliar language, we should strive to avoid being biased by our native language or other languages we know well. The practice of pre-modern linguists that described non-European languages in terms of Latin grammar has been thoroughly discredited. Now that English grammar has replaced Latin grammar as a tradition that is (almost) universally known among linguists, we do not want to repeat the errors of the pre-modern era and carry over concepts from English grammar to other languages. Likewise, we do not want to be biased by influential descriptions of other languages. Thus, linguists describing Australian languages do not want their descriptions to be Dyirbalocentric, despite the enormous influence of Dixon's (1972) description of Dyirbal. Since the advent of the Boasian approach in ethnography and structuralism (both European and American) in linguistics, it has been the goal of descriptivists to approach a language without prejudice and to do justice to its system, regardless of what systems other languages might have. We want to describe each language in its own terms.

Now my observation is that this goal of prejudice-free non-aprioristic description (or analysis) conflicts with the idea that a description should be based on a framework. It is well known that some frameworks have an English bias (cf. Van Valin 2005, who criticizes Chomskyan generative grammar in this regard; see also Van Valin, this volume). But even if it were possible to create a framework that avoids the bias of a particular language, the framework itself would constitute a bias, a set of prejudices with which a language is approached. A metalanguage by definition provides a pre-established set of expressions with a certain meaning, and by limiting ourselves to such a metalanguage, we would not be able to do justice to a language whose system does not correspond exactly to the concepts provided by the metalanguage. As has been argued at length by Croft (2001) (see also Dryer 1997; Haspelmath 2007; 2009*a*; Cristofaro 2008), grammatical categories and relations are language-specific, for all we know at the moment.

Of course, things could be simple. There could be a small set of innate grammatical categories and relations ("substantive universals") from which languages may choose, and a simple grammatical architecture linking the various components of the grammar ("formal universals"). It would be the linguists' task to determine the substantive and formal universals (in other words, universal grammar), and this would constitute the framework. Since it is innate, all languages must be describable within this framework. If this picture corresponded to the reality of languages, linguists' life would be easy and description could be based on a framework. However, all practicing linguists know that things are vastly more complicated. If a universal grammar, as envisioned in the Chomskyan tradition, exists, we are still very far from knowing what it is like. Almost every language presents us with new puzzles, with new categories and structures that do not fit into our frameworks. The idea that a single uniform framework could be designed that naturally accommodates all languages is totally utopian at the moment. So instead of fitting a language into the procrustean bed of an existing framework, we should liberate ourselves from the frameworks and describe languages in their own terms.

This has in fact been practiced widely by grammarians in the 20th century, especially by linguists working in the Boasian tradition of linguistic fieldwork or the traditions of European or American structuralism. Let us now look at two concrete examples of framework-free description.

14.3.2 First example: Tagalog basic sentence structure

Schachter and Otanes (1972: 59–85), still under the influence of American structuralism, describe Tagalog basic sentence structure in its own terms, and the result is a picture that is rather different from what is found in English (with which the authors contrast Tagalog). The basic pattern of Tagalog is not [$_{\text{sentence}}$ NP VP], but [$_{\text{sentence}}$ Predicate Topic]. There is a very rough correspondence between the Tagalog Topic and the English Subject NP, as can be seen in (1a). But the Topic may also correspond to the English Direct Object, as in (1b), or an English Prepositional Object, as in (1c). It is defined by its position (following the Predicate) and by its marking (Topic marker *ang*, used with non-pronominal, non-proper name Topics), not by its semantic role, which may be quite diverse.

(1) Tagalog (Schachter and Otanes 1972)

a. [*Gumising*]$_{\text{PRED}}$ [*ang bata*]$_{\text{TOP}}$. (p. 60)
awoke TOP child
'The child awoke.'

b. [*Sinulat ko*]$_{\text{PRED}}$ [*ang liham*]$_{\text{TOP}}$. (p. 60)
wrote I.CORE TOP letter
'I wrote the letter.'

c. [*Sinulatan ko*]$_{\text{PRED}}$ [*ang titser*]$_{\text{TOP}}$. (p. 60)
wrote I.CORE TOP teacher
'I wrote to the teacher.'

d. [*Artista*]$_{\text{PRED}}$ [*ang babae*]$_{\text{TOP}}$. (p. 61)
actress TOP woman
'The woman is an actress.'

e. [*Artista*]$_{\text{PRED}}$ [*ang nagluto ng pagkain*]$_{\text{TOP}}$. (p. 61)
actress TOP cooked CORE food
'The one who cooked some food is an actress.'

However, Topics have a semantic peculiarity that has no counterpart in English syntax: they must be definite. The main word of the Predicate is often a Verb, as in (1a–c), but it may also be a Noun, as in (1d–e) or an Adjective, so that calling the Predicate a "VP" would not make sense from the Tagalog point of view. Likewise, the main word of the Topic is often a Noun, as in (1a–d), but it can also be a Verb, as in (1e). While English needs a special Relative Clause construction (*the one who ...*) to make a referential expression corresponding to Tagalog *ang nagluto ng pagkain*, Tagalog can combine the Topic marker *ang* directly with the verb *nagluto*. Thus, even describing the Topic as a kind of "NP" would be very misleading, and Schachter and Otanes do not do this. Concepts from Latin and English grammar such as "subject", "NP", and "VP" play no role in their description of Tagalog. The terms "Predicate" and "Topic" are taken from the Western tradition, but they are given meanings that are specific to Tagalog (hence the capitalization of the terms here.)

14.3.3 Second example: German sentence-level word order

Since Drach (1937), descriptions of German word order have often posited a sentence schema for German that consists of at least five linear positions: Prefield, Left Bracket, Middlefield, Right Bracket, and Postfield. This way of describing German word order has come to be known as "field topology". Drach, a European structuralist, noted explicitly that his description was an attempt to "separate it from the ways of thinking of Latin grammar", he wanted to present German in a way that was founded in "the nature of the German language", and he urged that German be studied "without presuppositions, from outside", and "not through the Latin lens" (Drach 1937: §4, §16).

A recent summary of German field topology is found in Zifonun et al. (1997, 2: 1,498–1,505). In field topology, the verbal complex is the central element of the sentence. Its two elements in main declarative clauses constitute the Sentence Bracket: see the boldface elements in (2a–c).

(2) a. *Das Kind* ***hat*** *den Apfel heute* ***gegessen.***
the child has the apple today eaten
'The child ate the apple today.'

b. *Mutti* ***ruft*** *dich heute wahrscheinlich* ***an.***
mom calls you today probably up
'Mom will probably call you today.'

c. *Er* ***ist*** *dann natürlich* ***gerannt*** *wie ein Verrückter.*
he is then naturally run like a fool
'Then of course he ran like crazy.'

The finite verb (*hat, ruft, ist* in (2a–c)) is the Left Bracket, the non-finite verb (*gegessen, gerannt*) or the verb particle (*an*) is the Right Bracket. The position before the finite verb is called the Prefield, the position inside the bracket is called the Middlefield, and the position following the right bracket is called the Postfield. Thus, all German sentences follow the schema in (3).

(3) Prefield–Left Bracket–Middlefield–Right Bracket–Postfield

A whole range of generalizations can be formulated in terms of this schema:

(i) The elements of the verbal complex occur in the Left Bracket (finite verb) and in the Right Bracket (particle, nonfinite verb, in this order) in clauses without a subordinator.
(ii) The Prefield can only be filled by one single constituent (cf. (4a), where *das Kind* and *heute* are two constituents).
(iii) The Postfield can only be filled by clausal and other heavy consitituents (though in the spoken language this condition is often relaxed) (cf. (4b), which is only possible in the spoken language, and not generally considered correct).
(iv) In main declarative clauses, the Prefield and the Left Bracket have to be filled, as in (4c).
(v) In polar questions (and a few other specific sentence types), the Prefield is empty, as in (4d).
(vi) In clauses with a subordinator, the subordinator occurs in the Left Bracket position, the Prefield is empty and the entire verbal complex occurs in the Right Bracket (the order is particle, non-finite verb, finite verb), as in (4e).

(4) a. **Das Kind heute* ***hat*** *den Apfel* ***gegessen.***
the child today has the apple eaten
'The child today ate the apple.'

b. *??Das Kind* ***hat*** *den* ***Apfel*** *gegessen heute.*
the child has the apple eaten today
'The child ate the apple today.'

c. **Mutti dich heute wahrscheinlich* ***an-ruft.***
mom you today probably up-calls
'Mom will probably call you today.'

d. ***Ruft*** *Mutti dich heute* ***an***?
calls mom you today up
'Will mom call you today?'

e. … ***dass*** *Mutti sie gestern wahrscheinlich* ***an-gerufen hat***.
that mom her yesterday probably up-called has
'… that mom probably called him yesterday.'

These generalizations do not exhaust the word order rules of German, but other regularities mostly have to do with information structure. Crucially, grammatical relations such as "subject" and "object" (terms from Latin grammar) or constituents such as "VP" (a concept derived from English grammar) play no role in field topology.

14.3.4 Possible disadvantages

Two possible disadvantages of the framework-free approach to theoretical grammatical research are obvious and should be mentioned here. Both have to do with difficulty. Framework-free grammatical descriptions are more difficult to construct and more difficult to understand than descriptions built on familiar frameworks.

That creating a coherent, framework-free description of a language requires a major intellectual effort was recognized by the American structuralists, who typically assigned their doctoral students the task of describing a little-known language in its own terms. In the 19th century, when the need to create a new system of categories for each language had not yet been recognized and the framework of Latin grammar was thought to be universally applicable, description per se was rarely considered sufficiently demanding to give the author much scientific prestige. Similarly, in the generative tradition the description of (part of) a language in the generative framework is not considered sufficiently challenging; furthermore, dissertation authors are normally required to make novel proposals about the framework itself.

In addition, it is also easier to understand a grammatical description if it is written in a familiar framework. To understand the descriptions of Tagalog and German that we just saw requires the reader to first comprehend the novel notions of Topic, Prefield, Middlefield, etc. But such considerations are of course irrelevant form a scientific point of view and cannot be used to argue for framework-bound grammatical theory. If each language has its own categories, then it is simply wrong to carry over a category from one language to another language, and to use a framework that was created for one set of phenomena to describe another set of phenomena in a different language. If the correct approach involves greater effort, we have to make this effort.

In practice, however, the difficulties of framework-bound description can be significant, too. Descriptive frameworks have tended to grow in complexity over the last few decades, and mastering a complex framework puts a heavy burden on both the author and the reader. Since this effort is not creative in the same way as framework-free description is, many students of linguistics still find it easier (and professors find it easier to teach), but it binds many resources that are freed in the approach advocated here.

Moreover, the recognition that each language has its own categories does not mean that one cannot learn from other languages, because languages tend to exhibit great similarities in their categories and grammatical patterns. A linguist who has studied twenty (framework-free) grammatical descriptions of different languages will find the twenty-first language description fairly easy to follow, because there will be much that looks familiar from earlier descriptions. Because of the striking similarities between languages, it is often possible to use familiar transparent terminology (e.g., "Noun" for a word class denoting mostly people and things in English, and "Noun" for a semantically similar word class in Tagalog), rather than completely new or opaque terminology ("class B words"). The capitalization of language-specific grammatical terms helps the reader to remember that these are different categories (as with proper names; e.g., Mérida in Spain and Mérida in Venezuela are different cities).

Another objection that has sometimes been raised against framework-free descriptions is that they are "unconstrained". In the following section, I argue that the idea that frameworks should be restrictive is fundamentally mistaken.

14.4 Restrictive frameworks and their problems

14.4.1 Explanation by constrained description

One of the main attractions of descriptive frameworks has been the widespread idea that proposed frameworks are not just convenient metalanguages for the explicit, formal(ized) description of any language, but that frameworks are themselves explanatory. Such framework-based explanation is derived from the understanding of frameworks as restrictive: A framework is intended to allow the description of only those languages that actually occur. This idea, which has been prominent in Chomskyan generative linguistics since the 1960s and has been very influential in related approaches as well, is often expressed by its proponents in terms of a notion of descriptive power. Obviously a framework should be powerful enough to describe all languages, but in addition, in this view, it should not be too powerful (or "unconstrained") and allow the description of all sorts of languages that never occur. In other words, a descriptive framework should be able to describe all possible languages, but impossible languages should not be describable by it. This approach is reflected in the following quotation from Travis:

> The problem that the principles and parameters framework seeks to solve is: How can a grammatical system be flexible enough to account for language variation while at the same time be, to a large extent, restricted in order to account for the relative ease of language acquisition and the impossibility of certain language types? (Travis 1989: 263)

If descriptive frameworks were conceived of in the simplest terms, as metalanguages for precise description, they could not have any explanatory role. Notice that outside the field of linguistics, metalanguages do not seem to have the role of excluding impossible phenomena. Ordinary language can describe impossible things ("a rectangular triangle") and events ("the stone fell upward"); the language of arithmetic can describe impossible numbers ("33/0", or thirty-three divided by zero); and the language of heraldry can describe ill-formed coats of arms (e.g., the coat of arms of Samogitia is a sable bear on a gules field, which violates the rule of tincture that a color may not be placed on another color, only on a metal).

But in linguistics, especially generative linguistics, descriptive frameworks have been given an explanatory role. The descriptive framework of generative syntax has been equated with a theory of children's initial state in language acquisition, also called universal grammar (UG). "Universal grammar provides a genuine explanation of observed phenomena" (Chomsky 1988: 61–2), in the sense that only grammars consistent with UG can be acquired by learners and hence occur as adult languages. The fact that some logically possible languages do not occur is expressed in the lack of a possible description in the framework, and it is explained by the hypothesis that the framework reflects the child's innate knowledge of grammar. Thus, the idea that descriptive frameworks should be restrictive (should not be "too powerful", or "unconstrained", or should not "overgenerate") in order to be explanatory presupposes a fairly strong commitment to innateness.

In sections 14.4.2–4.5 we will see four examples of explanation by constrained description. Then in section 14.4.6 we will see that alternative explanations are available for these phenomena, so that there is no good reason to invoke restrictive frameworks.

14.4.2 First example: X-bar theory

A simple example that illustrates the idea of explanation by constrained description is X-bar theory. Phrase structure rules in human languages are quite diverse, as shown in (5a–c), but some logically possible phrase structure rules seem never to occur, as shown in (6a–c).

(5) a. NP → D [$_{N'}$ N PP] (e.g., *the* [*horse on the meadow*])
b. VP → Adv [$_{V'}$ V NP] (e.g., *often* [*eats white flowers*])
c. PP → Adv [$_{N'}$ P NP] (e.g., *right* [*under the tree*])

(6) a. NP → VP [$_{N'}$ Adv P]
b. VP → P [$_{P'}$ NP Adv]
c. PP → [$_{V'}$ P NP] V

Phrase structure rules of the traditional sort are thus too powerful and unconstrained, but "with the development of X-bar theory in the late 1960s, substantive constraints were placed on the form that [phrase structure rules] could take, constraints which expressed a particular set of empirical claims about what possible phrase structure arrangements can be found across languages" (McCloskey 1993: 497). X-bar theory, as it has been widely adopted since the 1980s and 1990s, basically only allows phrase structures of the type XP → YP [$_{X'}$ X ZP]. Other phrase structure rule types cannot be formulated in the X-bar framework and thus their non-existence is explained.

14.4.3 Second example: inflection outside derivation

Greenberg (1963, Universal 28) had observed that derivational affixes always come between the root and inflectional affixes when both inflection and derivation occur on the same side of the root. Anderson (1992) proposed a model of the architecture of universal grammar from which this generalization can be derived: If the lexicon and syntax are two separate components of grammar, and derivation is part of the lexicon, while inflection is part of the syntax, and if rules of the syntactic component, applying after lexical rules, can only add material peripherally, then Greenberg's generalization follows from the model of UG. Words with inflection inside derivation cannot be described in this model and thus their presumed non-existence is explained.

14.4.4 Third example: antisymmetry and word order asymmetries

Kayne (1994) discusses the mainstream view of phrase structure in generative grammar (i.e., X-bar theory) and finds it "overly permissive", "too unconstrained". He proposes that the precedence relation and the hierarchical relation of c-command should not be independent of each other but should be directly linked: "If X asymmetrically c-commands Y, x precedes y" (where X and Y are nonterminals and x and y are terminals they dominate; Kayne 1994: 33). This proposal (called *antisymmetry*) entails that all languages have an underlying SVO order, and other surface orders must be derived by movement. This has a number of interesting empirical consequences. For instance, in languages with clause-final complementizers, one has to assume that the entire rest of the clause ("IP") moves to a position preceding the complementizer (C) because underlyingly the complementizer (as the head of the clause) must be clause-initial. Thus, a sentence such as Japanese (7a) has the underlying structure (7b) and the derived structure (7c).

(7) Japanese

a. *Yooko-wa Masa-o aisite iru ka?*
Yoko-TOP Masa-ACC loving is Q
'Does Yoko love Masa?'

b. [$_{\text{CP}}$ [$_{\text{C}}$*ka*] [$_{\text{IP}}$ *Yooko-wa* [$_{\text{VP}}$ *aisite iru Masa-o*]]]

c. [$_{\text{CP}}$ [$_{\text{IP}}$ *Yooko-wa* [*Masa-o*$_i$ [$_{\text{VP}}$ *aisite iru* t$_i$]]]$_j$ [$_{\text{C}}$*ka*] t$_j$]

d. *Yooko-wa dare-o aisite iru ka?*
Yoko-TOP who-ACC loving is Q
'Whom does Yoko love?'

The landing site for this movement is presumably the specifier of C, a position that in many languages is used as a landing site for *wh*-movement in questions. According to Kayne (1994: 54), this explains that OV languages with final complementizers like Japanese tend not to have *wh*-movement in questions, as shown in (7d). In Kayne's antisymmetry framework, such languages cannot be described and thus their non-existence is explained.

14.4.5 Fourth example: argument-flagging in Optimality Theory

Like other brands of generative grammar, mainstream Optimality Theory (OT) practices explanation by constrained description. According to McCarthy,

> One of the most compelling features of OT, in my view, is the way that it unites description of individual languages with explanation in language typology ... OT is inherently typological: the grammar of one language inevitably incorporates claims about the grammars of all languages. (McCarthy 2002: 1)

A striking difference between OT and the proposals in the preceding subsections (14.4.2–4.4) is that the interesting aspects of the framework are the constraints, which are often fairly concrete, and not highly abstract principles such as antisymmetry or the lexicon-syntax bifurcation. There is thus often a more direct relationship between the explanatory mechanisms (the constraints) and the explananda (the cross-linguistic patterns).

Here I have space only for one concrete example, the distribution of argument-flagging patterns (i.e., case and adpositional marking) in intransitive and transitive clauses, as discussed and explained by Woolford (2001). Woolford observes that languages may show the patterns in (8a) but do not generally show the patterns in (8b).

(8)		intransitive patterns	transitive patterns
	a.	nominative	nominative–accusative
		ergative	ergative–nominative
		dative	dative–nominative
	b.	accusative	ergative–accusative
			dative–accusative

Woodford explains these patterns by positing for each of the cases a markedness constraint against it, and a universally fixed ranking of these constraints: *ERGATIVE/*DATIVE≫*ACCUSATIVE ≫*NOMINATIVE. This means that other things being equal, nominative is favored over accusative and accusative is favored over ergative and dative (ergative and dative are not ranked with respect to each other). In addition, Woolford posits a faithfulness constraint FAITHLEX, which requires that the lexically specified case features must appear on the surface. (The presupposition is that agentive arguments are lexically specified as [+ergative subject], and experiencer subject arguments as [+dative subject].)

Given this system, languages that do not allow non-nominative subjects at all (such as English) are described by the ranking *ERGATIVE/*DATIVE ≫ FAITHLEX ≫*ACCUSATIVE ≫*NOMINATIVE, i.e., in these languages the markedness constraints against ergative and dative outrank faithfulness. Since nominative is universally least marked, it appears instead of ergative or dative. In languages where faithfulness to role-based lexical specification is ranked higher, ergative and dative subjects can surface (as in Basque and Japanese, for instance). Crucially, the object of ergative/dative subject clauses can never appear in the accusative because accusative is less favored than nominative. The intransitive argument cannot appear as accusative for the same reason: *ACCUSATIVE is universally ranked higher than *NOMINATIVE, so that the nominative candidate always wins the day. Accusative case appears only when another nominative is present because a higher constraint against equal cases in transitive clauses rules out the nominative–nominative pattern.[2] Thus, a language with intransitive accusative arguments or transitive ergative–accusative or dative–accusative patterns cannot be described in this system, while attested language types can be described by different constraint rankings.

Analyses of various split marking patterns have been proposed by Aissen (1999; 2003) in much the same spirit as Woolford's. I have discussed and criticized Aissen's proposals elsewhere (Haspelmath 2008*e*; 2009*e*).

[2] Woolford assumes another constraint, which is unviolable and outside her OT analysis, that restricts accusative to positions within VP, thus accounting for the impossibility of the accusative–nominative pattern.

14.4.6 Against restrictive frameworks and explanation by constrained description

As we saw in section 14.4.1, the general strategy of explaining observed constraints on attested languages by a constrained descriptive apparatus presupposes the assumption that this descriptive apparatus is innate (i.e., the assumption of universal grammar). The basic idea is that unattested languages are unacquirable languages. For some reason, generative linguists have by and large ignored the possibility of constraints on attested languages coming from factors of language use rather than language acquisition. But if explanations from language use (also called *functional explanations*) are considered seriously, it soon becomes apparent that they can account for a wide range of constraints on attested languages (cf. Moravcsik 2009). To be transmitted in a speech community, a language must be usable, not just acquirable. This point has occasionally even been made by generative linguists (see the quotations below), but its consequences for the enterprise of framework-bound grammatical theory have not been widely realized.

> [T]he scope of the language faculty cannot be derived even from an exhaustive enumeration of the properties of existing languages, because these contingent facts result from the *interaction* of the language faculty with a variety of other factors, including the mechanism of historical change . . . [O]bservations about preferences, tendencies, and which of a range of structural possibilities speakers will tend to use in a given situation are largely irrelevant to an understanding of what those possibilities are. (Anderson 1999: 121)

> [M]any of the so-called *phonological universals* (often discussed under the rubric of markedness) are in fact epiphenomena deriving from the interaction of extragrammatical factors like acoustic salience and the nature of language change . . . Phonology [i.e., a theory of UG in this domain, M.H.] is not and should not be grounded in phonetics since the facts that phonetic grounding is meant to explain can be derived without reference to *phonology*. (Hale and Reiss 2000: 162)

> It is not the job of generative theory to account for typological generalizations. Attempts to do so by means of parameterized principles have been failures. Such generalizations belong to the domain of performance, rather than to the domain of formal grammar and, as a consequence, Universal Grammar itself can be relieved of the responsibility of accounting for them. (Newmeyer 2005: 126–7)

In Haspelmath (2004*a*), I have summarized the arguments against basing a theory of the cognitive code for language (= universal grammar) on the range of attested languages, pointing out that the situation in biology is quite parallel: The genetic code allows a much wider range of organisms than are actually found in nature. The narrow range of actually existing organisms is primarily determined by survival (i.e., the chance of successful replication), not by constraints on what the genetic code allows. To study the nature of the cognitive code, we should study the acquisition of unattested language types under natural or artificial conditions, but we should not hope to derive much insight from constraints on attested languages.

Most of these constraints have very good functional explanations, i.e., explanations deriving from different chances of being replicated in language use (Croft 2000).

For instance, the major true generalizations of X-bar theory (section 14.4.2), that phrases of particular types have heads of particular types, can easily be explained by the task of syntax to express conceptual consituents with similar conceptual structures (cf. Jackendoff 1983; 2002, Chapter 12). Attempts at extending X-bar theory from NPs, VPs, and PPs to other syntactic phrases (such as IP, CP, FocP) are not particularly plausible and have not been fruitful outside a particular narrow framework.

Another example is the position of inflectional affixes and derivational affixes with respect to each other and to the stem (section 14.4.3). Bybee (1985*a*: 33–5; 1985*b*) has shown that there is a broader generalization such that grammatical categories whose meaning is more relevant to the verb stem's meaning tend to occur close to it, subsuming Greenberg's Universal 28 under it. She attributes this regularity to iconicity: Meanings that are more relevant to each other are mirrored by forms that occur closer to each other.

Next, what about the position of *wh*-phrases in a clause and other word order properties of the language (section 14.4.4)? Hawkins (2002, §4.3; 2004*a*, §7.3) argues that *wh*-movement creates filler-gap relationships that cause processing difficulty and that the processing difficulty is greater if the verb (to which most *wh*-phrases are connected semantically) is further away. This predicts that VSO languages should favor *wh*-fronting the most, while SOV languages should favor it the least, with SVO languages in between, and this is borne out by the available cross-linguistic data.[3]

And finally, the occurrence of various argument-flagging patterns in transitive and intransitive clauses is also amenable to a functional explanation. With core arguments, the most important role of argument flagging is distinguishing the arguments, and for this it is sufficient if one of them is marked overtly. The case that is not marked overtly is generally called "nominative", so this functional consideration is sufficient to explain the absence of ergative–accusative and dative–accusative patterns. It does not explain an alleged asymmetry that Woolford's OT system captures: According to Woolford, intransitive clauses with a single ergative argument occur (e.g., in Basque), but intransitive clauses with a single accusative argument do not occur. However, this claim is not backed up with cross-linguistic data, and it is not difficult to find in the literature examples of languages whose intransitive clauses may have accusative single arguments. A language of this kind (the mirror image of Basque) is Central Pomo (a language of California; Mithun 1991: 518–23):

[3] In the data of Dryer (2005*b*) and (2005*c*), the figures are as follows (the figures refer to languages, before the slash, and genera, after the slash):

	SOV	SVO	VSO
wh-fronting	52/38	65/35	42/23
no *wh*-fronting	225/109	188/57	16/6

(9) a. *ʔa·* *mú·ṭ̯u* *ʔé·yčadiw.* (p. 518)
I.NOM he.ACC chased. away
'I chased him away.'

b. *Mu·l* *qʰaʔán·ṭ̯aw.* (p. 522)
he.NOM dreamed
'He was dreaming.'

c. *Q'alá·w* *mú·ṭ̯u.* (p. 521)
died he.ACC
'He died.'

This is not common, but languages like Basque in which some intransitive single arguments may be in the ergative are not common either, so it is not clear that there is a generalization that needs to be explained.

Thus, gaps in the observed range of linguistic diversity typically have functional explanations, and there is no need to invoke theoretical frameworks (reflecting the innate universal grammar) to explain them (this point is also made by Dryer 2006*a*, using similar arguments).

But innate theoretical frameworks are not only unnecessary, they are also insufficient to explain gaps in typological patterns. The reason is that framework-based explanation can only explain absolute universals, but not statistical universals (or universal tendencies). However, most empirical universals are tendencies. There are numerous exceptions to the generalization that inflection occurs outside derivation (e.g., Bochner 1984, Rainer 1996), numerous exceptions to the generalization that languages with final subordinators do not have *wh*-fronting (the databases of Dryer 2005*c* and 2005*a* contain 33 such languages), and, as we just saw, there are exceptions to the generalization that intransitive clauses with a single accusative argument do not occur.

Another serious problem with framework-based/UG-based explanation of typological patterns is the diversity of categories across languages. Strictly speaking, categories such as "accusative case", "inflection", and "preposition" cannot be defined across languages but only in language-specific terms (Dryer 1997; Croft 2001; Haspelmath 2007). This means that it is unclear how the claims made by innatist frameworks should be tested. Proponents of framework-based description and explanation tend to simply ignore this problem.

I conclude that a major reason for adopting universally applicable descriptive frameworks in theoretical linguistics is not well founded: Frameworks, interpreted as innate restrictions on what can be acquired, are not well suited to explaining patterns in linguistic diversity. But descriptive frameworks have also been proposed by functional linguists with little or no interest in the generative enterprise of explanation by constrained description, so we should now turn to such functional frameworks.

14.5 Functional frameworks and their problems

The two most prominent frameworks developed by functional linguists are Functional Grammar (FG, see Dik 1997)[4] and Role and Reference Grammar (RRG, see Van Valin 2005, Van Valin 2009). Since other functionalist approaches are framework-free and do not propose a universally applicable set of concepts for structure description, these two frameworks are sometimes called "structural-functional theories" (e.g., by Butler 2003, who provides a detailed comparative discussion of FG, RRG, and Michael Halliday's Systemic Functional Grammar). Linguists working in these frameworks do not assume that the framework's concepts and structures are innate, and they do not try to explain gaps in attested languages by making the framework restrictive. So in the practice of these linguists, there is no place for explanation by constrained description, but what they share with generative linguists is the assumption that there is a set of universal categories and concepts by which all languages can be described in an insightful way. These frameworks are thus as aprioristic as generative grammar, and they inherit the problems of apriorism. Both FG and RRG emphasize that they want to avoid the well-known Anglocentrism of generative syntax, but they do not draw the conclusion (which I regard as compelling) that one should not approach languages with a pre-established set of concepts at all and describe each language in its own terms, i.e., without a framework. Van Valin (2005: 1) asks: "What would linguistic theory look like if it were based on the analysis of languages with diverse structures such as Lakhota, Tagalog and Dyirbal, rather than on the analysis of English?" This describes precisely the problem that a non-aprioristic, framework-free approach tries to avoid: The analysis of one language should never be "based on" the analysis of another language. Lakhotacentric or Tagalocentric frameworks are in no way better than Anglocentric frameworks.

Let me illustrate some concrete problems arising from apriorism in FG and RRG, using the example of ditransitive constructions (cf. also Haspelmath 2008*c*). A much-discussed issue is the description of contrasts such as that between the Prepositional Dative Construction and the Double Object Construction in English:

(10) a. *Aisha gave the money to Pedro*.
     b. *Aisha gave Pedro the money*.

In FG, this is analyzed by saying that the recipient (*Pedro*) has the syntactic function of "object" in (10b) but not in (10a), where it is marked by the preposition *to* according to its semantic role, and where the theme (*the money*) has the object function

[4] Functional Grammar has meanwhile been superseded by Functional Discourse Grammar (see Hengeveld and Mackenzie 2008, 2009).

(Dik 1997, Chapter 10). In FG, "subject" and "object" functions are assigned only if there is an alternation, i.e., a passive construction or a "dative shift" construction as in (10a–b). Similarly, in RRG it is claimed that recipient (*Pedro*) is assigned the semantic macrorole of "undergoer" in (10b) but not in (10a), where the theme (*the money*) is assigned the undergoer role (Van Valin 2005: 114), as a "marked option". Both FG and RRG assume the universality (or at least cross-linguistic applicability) of their concepts "object" and "undergoer", and this leads to problems with languages that diverge from the English pattern in (10a–b). Many languages have only a pattern that resembles (10b) but no pattern resembling (10a). In FG, this would mean that object assignment is obligatory, counter to a principle of the theory (cf. Dik 1997: 282–5 for discussion), and, in RRG, it would mean that a language has "marked" undergoer assignment as the only option, counter to the spirit of markedness (cf. Van Valin 2005: 123–7 for discussion). Van Valin eventually revised his principles for actor and undergoer selection in a fairly drastic way in recognition of this, leading to a more complex, less elegant descriptive theory (Van Valin 2007).

Thus, although both FG and RRG have always been aware of the problems of potential Anglocentrism, they were not able to avoid an Anglocentric proposal for this particular phenomenon, presumably because at the time when the proposals were first made (around 1980), no significant cross-linguistic research on ditransitive constructions had been carried out. So one lesson is that it seems to be impossible to construct a non-biased framework unless one has done a significant amount of cross-linguistic research. But cross-linguistic research is always preliminary, and thus the framework is always biased against those languages that have not been studied yet. And a second lesson is that frameworks that can extend to more languages equally naturally are inevitably more complex and less elegant. The question is how complex the framework will be once the full range of cross-linguistic evidence has been examined. My suspicion is that it will be so complex that it is not really distinguishable anymore from the position advocated here, i.e., not to work with a catch-all framework but to construct the needed descriptive categories anew for each language.

14.6 Basic Linguistic Theory and its problems

Some authors (notably Dixon 1997: 128–38 and Dryer 2006*b*) have emphasized that descriptive work on the world's languages resulting in reference grammars is by no means "merely descriptive", but is theoretical, not just in the general sense (sense 4 of section 14.2.3) but also in the sense of "theoretical framework". These authors refer to the theoretical framework employed by grammar writers, historical linguists, and typologists as *Basic Linguistic Theory*. They would probably object to the main thrust of this chapter and argue that grammatical theorizing should not be framework-free, but should use the framework of Basic Linguistic Theory. Dryer (2006*b*), in particular, notes that frameworks in the Chomskyan tradition are intended as descriptive and explanatory theories at the same time, and argues that if one drops the nativist presuppositions of this approach, then one must conclude that languages are best described (and cross-linguistic generalizations, the basis for functional explanations, are best formulated) in terms of Basic Linguistic Theory.

However, Dixon and Dryer seem to contradict themselves when they emphasize that work in the framework of Basic Linguistic Theory attempts to describe languages in their own terms rather than on the model of a well-known language or of some prestigious framework. According to Dixon,

> When writing a grammar in terms of Basic Linguistic Theory one takes nothing for granted. Each analytic decision has to be approached as an open question . . . In contrast, each of the non-basic theories posits that certain categories are relevant for all languages–one only has to find them. (Dixon 1997: 132)

Similarly, Dryer observes that

> Basic Linguistic Theory differs from traditional grammar most strikingly in its attempt to describe each language in its own terms, rather than trying to force the language into a model based on European languages. (Dryer 2006*b*: 211)

The contradiction lies in the claim that "one takes nothing for granted" and each language should be described "in its own terms", while at the same time it is claimed that Basic Linguistic Theory consists of certain concepts that grammar writers must know before they can describe a language ("the fundamental theoretical concepts that underlie all work in language description", Dixon 1997: 128; "the descriptive tools assumed in descriptive grammars", Dryer 2006*b*: 210). What Dixon and Dryer probably have in mind when they refer to "theoretical concepts" or "descriptive tools" of Basic Linguistic Theory is the kinds of concepts that are presented in works such as Payne (1997) and Shopen (2007), two widely used works that prospective grammar authors are typically directed to for inspiration.

However, if these concepts and tools are treated as a true framework, i.e., as a set of options from which descriptivists and languages may choose, they defeat the stated goal of open-minded, bias-free description. Grammar authors have to be ready to create completely novel concepts, because no two categories are completely identical across languages, and often the categories are not even particularly similar across languages. If one approaches a language with a particular set of concepts and tools in mind, one is no longer open-minded and bias-free.[5]

[5] Matthew Dryer (p.c.) has told me that he regards the principle of describing each language in its own terms as the most important principle of Basic Linguistic Theory. If this is so, Basic Linguistic

I hasten to add that the kinds of concepts found in typologically oriented handbooks for grammar writers (such as Payne 1997 and Shopen 2007) are very useful to know for every linguist and that, by making use of these concepts, grammar writers will probably write less biased grammars than if they use other frameworks. But it remains true that, ideally, they would not make use of pre-established concepts and tools but would create the tools they need during the process of writing the grammar.

Fortunately, in actual fact, this is what grammar writers do most of the time or in any event very often and characteristically. They introduce concepts that are justified by the phenomena of the language at hand and that need no justification beyond it. They do not feel bound by a particular framework, but they create new concepts as they see the need for them.

Thus, I do not accept the assertion that "there is no such thing as atheoretical description" (Bach 2004: 50; Dryer 2006*b*: 207), if "atheoretical" here means "framework-free" (as it seems to mean from the context). I agree with Dixon (1997: 134) that "every person who describes a language is also a theoretician . . . Every point in a grammatical description is a theoretical statement, and must be justified by appropriate argumentation" (if "theoretician" is meant in sense 3 or 4 of "theory"; see section 14.2.3), and also with Dryer (2006*b*: 212) that "the analytical concepts one assumes necessarily constitute a set of theoretical assumptions", but one can make theoretical statements without presuppositions about which concepts should be used.[6]

Dixon implies that his own work is formulated in terms of Basic Linguistic Theory, but, on closer examination, his work is full of concepts that are by no means readily applicable to any language. Consider one of the examples he mentions in Dixon (1997: 132): "Is it appropriate to recognise one unit 'word' or two (a 'phonological word' and also a 'grammatical word')?" Dixon's view that phonological and grammatical words may but need not coincide is well known (see also Dixon and Aikhenvald 2002), but he does not seem to allow for the possibility that languages do not make use of a word-like unit at all, or make use of several different phonological and grammatical words, or make use of a word-like unit that is defined by both phonological and grammatical criteria but contrasts with other word-like units. The framework-free approach allows for these possibilities as well.

Theory would be equivalent to framework-free grammatical theory as advocated here, and it could not be a "descriptive/theoretical framework" in the sense of this chapter.

[6] I would be happy to accept the possible view (which I have not seen expressed by anyone) that a description of a language necessarily involves a framework, but that it could (and should) be a different framework for each language. This would be equivalent to what I am proposing, but since the term *framework* has always been used for universally applicable frameworks, I chose to argue here against frameworks *tout court* rather than against "universally applicable frameworks". This is of course just a terminological matter.

In some of his works Dixon insists on a particular meaning of a traditional term, as when he emphasizes that a *predicate* in linguistics is a verb and its modifiers, not (as in Greek logic) what remains of a clause after substracting the subject, so that a copula complement should not be called "predicate nominal" (Dixon 2004: 7). There is nothing wrong with such a terminological choice, but it is misleading to suggest that Dixon's proposals are equal to "the fundamental theoretical concepts that underlie all work in language description" (= his definition of Basic Linguistic Theory). Work in language description operates with a wide range of theoretical concepts, and with a fair amount of terminological diversity. But it tends to be terminologically conservative, and this seems to have led to the view that the concepts used in language description are also conservative (cf. Dryer's (2006*b*: 211) assertion that Basic Linguistic Theory can be "roughly described as traditional grammar, minus its bad features"). But this is not necessarily the case. Good descriptive grammars do not adopt their concepts from earlier work, but they are often terminologically conservative because they want to reach a wide audience (unlike works in particular frameworks, which mostly address colleagues working within the same framework and can therefore be terminologically innovative).

14.7 Framework-free comparative and explanatory theory

Since all languages have a huge amount of properties that are due to historical accidents and cannot be explained except with reference to these accidents, true explanation in linguistics is restricted to explanation of language universals. Explanatory theoretical work must therefore adopt a broadly comparative approach, a point about which there is widespread agreement:

> In order to explain the data in individual languages, a theory must make falsifiable empirical claims about the entire class of natural languages. (Perlmutter 1980: 196)

> The generativist will have to compare English with other languages to discover to what extent the properties he has identified are universal and to what extent they are language-specific choices determined by universal grammar . . . Work in generative linguistics is therefore by definition comparative. (Haegeman 1994: 18)

In Chomskyan generative linguistics, the descriptive framework also plays a crucial role in comparison and explanation. As we saw, it is assumed that the same framework can be applied to all languages, and that once the right framework has been found, it can also be used to compare the languages, in order to determine how they differ. This is a very difficult process, because the framework is both the ultimate result of the comparison (a characterization of UG explaining the limits on variation) and a prerequisite to the comparison (languages cannot be compared unless they are first described in the correct framework, Newmeyer 1998: 337–8). As a result, comparative studies in the generative framework have not been very successful, at least much less successful than was expected in the 1980s, when the Principles and Parameters program was initiated (Haspelmath 2008*b*).

By contrast, framework-free comparative linguistics is thriving (e.g., Haspelmath et al. 2005). Large-scale cross-linguistic comparison without a framework is not free of difficulties either, but it has become easier because of the availability of a steadily increasing number of detailed reference grammars written in a framework-free but accessible format. Dixon (1997: 128, 132) has claimed that Basic Linguistic Theory is the framework that underlies such typological work, but this is not correct. Typological work as represented by *The World Atlas of Language Structures* (WALS) is just as framework-free as most of the grammatical descriptions it is based on, though it is of course highly theoretical, just like the descriptive work it depends on.

In this regard, Dixon's view of the role of Basic Linguistic Theory in linguistics is similar to the generative view: The same concepts are used for description and comparison. However, in actual typological practice, a rather different picture emerges. Typologists make up their own concepts (called *comparative concepts* in Haspelmath 2009*a*) and match them against the facts of each language, but they do not expect to find the same categories in all languages, and their comparisons can accommodate great variation in the actual categories of languages (called *descriptive categories* in Haspelmath 2009*a*). For instance, typologists often work with a comparative concept "ergative case" (overt case of the transitive agent as opposed to the case of the intransitive single argument), but if a language has a case that marks both the transitive agent and the possessor (like the Eskimo Relative case), this also counts as an ergative case. Cases that count as ergative can thus be quite diverse. Similarly, typologists work with the comparative concept of "adjective" (= property word), but if a language has a word class ("Verb") comprising both action words and property words, they still count as adjectives in the comparative sense. Again, words that count as adjectives can be very diverse. As a final example, consider the comparative concept "*wh*-word" (used to question particular sentence parts). If a language has a class of "indeterminate" pronouns that can be used both for questioning and for indefinite reference ("who; someone"), these count as *wh*-words, too. Thus, the typologists' comparative concepts are not necessarily equatable with the descriptive categories of languages.

Since grammatical categories are different in different languages (just as word meanings are different in different languages), comparative linguists cannot help but create specific concepts for the purpose of comparison (comparative concepts). The criterion of adequacy for comparative concepts is not the correctness of the description (as for descriptive categories), but the fruitfulness of the resulting comparison (see Haspelmath 2009*a*). Since comparativists can approach languages from multiple angles, there is no single correct set of comparative concepts. In WALS, for example, different authors have opted for slightly different "case" concepts (Baerman and Brown 2005, Iggesen 2005), but there is no contradiction. The concepts are not identical, only the chosen terms happen to coincide. Like descriptive grammarians, typologists tend to be terminologically conservative because their work addresses a wide range of potential users. This practice should not be mistaken as the use of a common framework by all typologists.

In the approach advocated here, explanatory theory primarily consists of functional explanation (cf. section 14.4.6 above). Like functional explanation in biology (cf. Nettle 1999), functional explanation in linguistics is necessarily diachronic (Bybee 1988*a*; Keller 1994; Kirby 1999; Haspelmath 1999*b*; 2008*a*). As Dryer (2006*a*: 56) puts it, "a theory of why languages are the way they are is fundamentally a theory of language change". Explanatory grammatical theory of this sort (as exemplified by works such as Givón 1979; Bybee et al. 1994; Heine 1997; Frajzyngier and Shay 2003; Hawkins 2004*a*) has no need for (descriptive/theoretical) frameworks.

14.8 Conclusion

This is not the first work to reject framework-bound grammatical theorizing. Lazard (2006: 93) says: "Le descripteur doit se garder de tout modèle" ("Descriptivists should beware of any model/framework"). And Givón said:

> "Framework", "format", "theory" and "Grammar" are words that have been much maligned in the past three decades in linguistics. Ever since the Bloomfieldians, such labels have meant, more likely than not, the closing of one's horizons and the wedding of oneself to a restrictive, counter-empirical and anti-explanatory formalism. (Givón 1984: 25)

Even though Givón did not include this statement in the revised (2001) version of his two-volume work on syntax, I still think that he was basically right in 1984.[7] When approaching a language, we should not close our horizons by applying an aprioristic, pre-established framework to it.

I have argued here that the set of concepts needed for the description (or analysis) of a language must be constructed separately for each language because all languages have different structures. I gave two extended examples from well-known framework-free descriptions of Tagalog and German clause structure, and I noted

[7] In Givón (1995), there are two chapters entitled "Taking structure seriously", in which Givón tries to counter a perceived "grammar denial syndrome" among some functionalists. In view of the flood of descriptive grammars that have been written in the last two decades, I see no sign of such a trend (except perhaps among a few American functionalists who shifted their interests from grammar to discourse).

that many good grammars follow these examples, even though the originality of the descriptions (or analyses) is often concealed by the use of familiar terminology. I observed that in generative linguistics, frameworks are invoked both for description and for explanation (by constrained description), and that the idea that frameworks should be restrictive makes sense only if they are equated with an innate universal grammar. I further noted that structural-functional descriptive frameworks and the descriptive framework of Basic Linguistic Theory also contradict the methodological imperative of bias-free grammatical analysis, and that explanatory theory does not consist in the construction of frameworks but in (ultimately diachronic) functional explanation of universal tendencies.

At this point, some readers may ask: If there are no frameworks, then what should I teach my students in syntax classes? My answer is: The best syntax class is a field methods course, and the second best syntax class is a typology course. If we want to understand the nature of syntax, we have to study the syntactic patterns of concrete languages, preferably unfamiliar languages, to broaden our horizons. Since they cannot get first-hand experience of a larger number of languages, students should study existing framework-free descriptions of languages from around the world, be encouraged to ask critical questions about each analysis, and learn to compare languages with diverse categories by means of universally applicable comparative concepts.

This page intentionally left blank

CHAPTER 15

FUNCTIONAL DISCOURSE GRAMMAR

KEES HENGEVELD
J. LACHLAN MACKENZIE

15.1 Introduction[1]

This chapter introduces Functional Discourse Grammar (FDG), a typologically-based model of language structure. After a general outline of the model and its place as the grammatical component of a wider theory of verbal interaction in section 15.2, section 15.3 will situate the model within the field of grammatical theories at large. Section 15.4 will deal with the details of the four levels of linguistic organization (interpersonal, representational, morphosyntactic, and phonological) inside the grammar proper, giving examples of the potential of each. Section 15.5 will give an impression of how both the interaction of the grammar with surrounding components and the interaction between the various levels within the grammatical component help explain a wide range of linguistic phenomena. After a detailed analysis of a worked example in section 15.6, we will discuss some further applications of FDG in section 15.7.

[1] For a full account of FDG see Hengeveld and Mackenzie (2008). Mackenzie's contribution to this article was partially funded by the Spanish Ministry of Science and Innovation, HUM2007-62220.

15.2 Outline of the model

15.2.1 FDG and verbal interaction

As shown in Figure 15.1, FDG is conceived of as the Grammatical Component of an overall model of verbal interaction in which it is linked to a Conceptual Component, an Output Component, and a Contextual Component. These three non-grammatical components interact in various ways with the Grammatical Component, more specifically with the operations of Formulation and Encoding. Formulation concerns the rules that determine what constitute valid underlying pragmatic and semantic representations in a language. Encoding concerns the rules that convert these pragmatic and semantic representations into morphosyntactic and phonological ones. FDG assumes that both Formulation and Encoding are language-specific, i.e., no universal pragmatic, semantic, morphosyntactic, or phonological categories are postulated until their universality has been demonstrated through empirical research.

The Conceptual Component is responsible for the development of both a communicative intention relevant for the current speech event and the associated conceptualizations with respect to relevant extra-linguistic events, and is thus

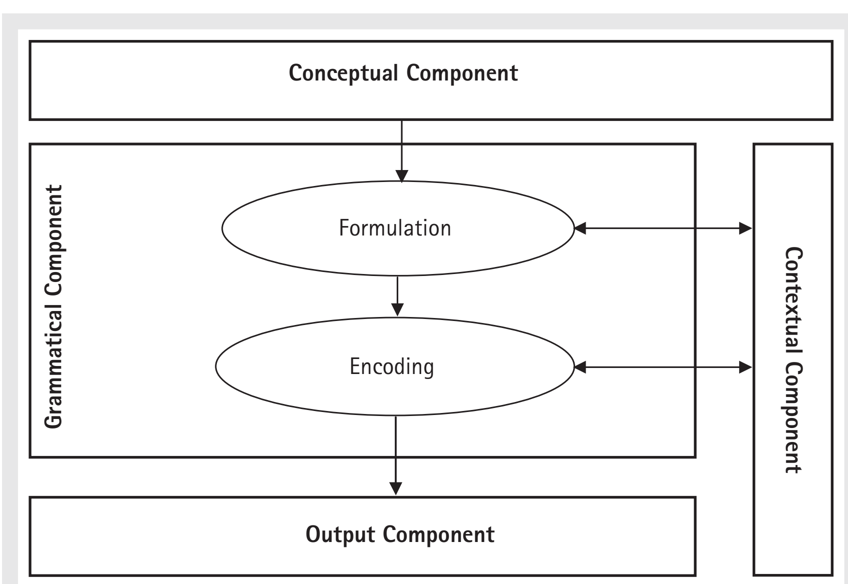


Figure 15.1. FDG as part of a wider theory of verbal interaction

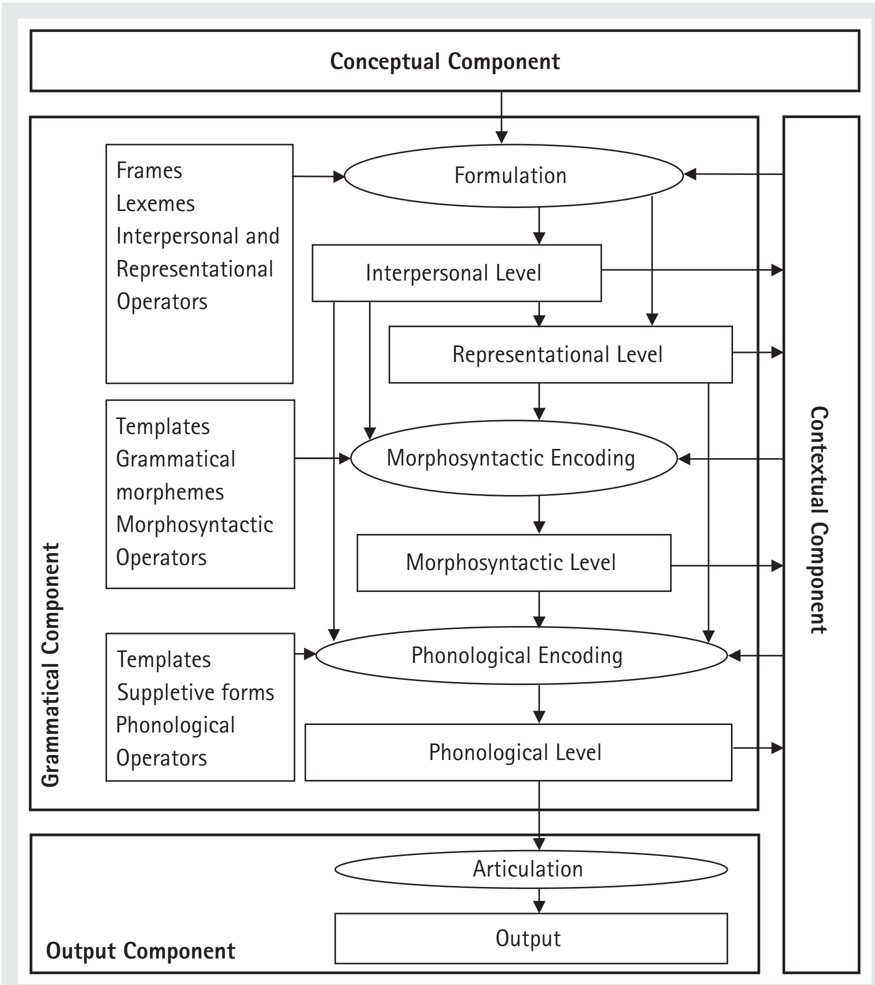


Figure 15.2. General layout of FDG

the driving force behind the Grammatical Component as a whole. The Output Component generates acoustic or signed expressions on the basis of information provided by the Grammatical Component. Its function may be seen as translating the digital (i.e., categorical, opposition-based) information in the grammar into analog (i.e., continuously variable) form. The Contextual Component contains a description of the content and form of the preceding discourse, of the actual perceivable setting in which the speech event takes place, and of the social relationships between Participants. This type of information is relevant to many grammatical processes, such as narrative chaining, reflexives, and passives.

15.2.2 The architecture of FDG

The general architecture of FDG itself, in relation to the components that flank it, may now be represented as in Figure 15.2, in which the Grammatical Component is presented in the center, the Conceptual Component at the top, the Output Component at the bottom, and the Contextual Component to the right.

A distinguishing feature of FDG shown in Figure 15.2 is its rigorous top-down architecture: FDG starts with the speaker's intention and works down to articulation. This is motivated by the assumption that a model of grammar will be more effective the more its organization resembles language processing in the individual. Psycholinguistic studies (e.g., Levelt 1989) clearly show that language production is indeed a top-down process. The implementation of FDG reflects this process and is organized accordingly. This does not mean, however, that FDG is a model of the speaker: FDG is a theory about grammar, but one that tries to reflect psycholinguistic evidence in its basic architecture. The top-down organization of the grammar has far-reaching consequences at all levels of analysis, as we will show in section 15.4.

In Figure 15.2 ovals contain operations, boxes contain the primitives used in operations, and rectangles contain the levels of representation produced by operations. We will discuss all of these in more detail in section 15.4, and here limit ourselves to describing the general top-down process on the basis of a simple example, given in (1), produced in a context in which the Addressee wants to enter a field that hosts a bull:

(1) There's a bull in the field!

In the prelinguistic Conceptual Component a communicative intention (issuing a warning) and the corresponding mental representations (of the event causing danger) are relevant. The operation of Formulation translates these conceptual representations into pragmatic and semantic representations at the Interpersonal and the Representational Levels respectively. Warnings are not a separate illocutionary category in English, but the Speaker solves this problem by selecting a Declarative Illocution combined with an Emphatic operator at the Interpersonal Level. The entity causing danger is furthermore characterized as a Focal Topic at this Level. At the Representational Level the Speaker chooses to designate the entity causing danger as part of a locative predication frame. The configurations at the Interpersonal and the Representational Levels are translated into a morphosyntactic structure at the Morphosyntactic Level through the operation of Morphosyntactic Encoding. In (1) this involves, for instance, the word order characteristic of existentials, the insertion of dummy *there*, etc. Similarly, the structures at the Interpersonal, Representational, and Morphosyntactic Levels are translated into a phonological structure at the Phonological Level. In this case, for instance, the selection of the Declarative Illocution combined with an Emphatic operator is responsible for the overall intonation contour with a high fall on the Focal Topic *bull*. By organizing the Grammatical Component in the way illustrated here, FDG takes the functional approach to language to its logical extreme: within the top-down organization of the grammar, pragmatics governs semantics, pragmatics and semantics govern morphosyntax, and pragmatics, semantics, and morphosyntax govern phonology.

The Phonological Level of representation is the input to the operation of Articulation, which contains the phonetic rules necessary for an adequate utterance. Articulation takes place in the Output Component, outside the grammar proper.

The various levels of representation within the grammar feed into the Contextual Component, thus enabling subsequent reference to the various kinds of entity relevant at each of these levels once they are introduced into the discourse. The Contextual Component feeds into the operations of Formulation and Encoding, so that, for instance, the availability of antecedents may influence the composition of (subsequent) Discourse Acts. Having seen something of the architecture of FDG let us now place it in its broader context.

15.3 Theoretical background

The main goal of Functional Discourse Grammar is to give an account of morphosyntactically and phonologically codified phenomena in languages, either as correlated with pragmatic or semantic aspects of Formulation or as displaying inherent properties of Encoding. In the former case, the phenomenon is functionally motivated; in the latter case, it is arbitrary. As the name of the theory suggests, the emphasis in FDG work is strongly on the former. The functionalist stance entails the hypothesis that a wide range of formal categories can be insightfully explained if they are brought into correspondence with semantic and pragmatic categories rooted in human cognition and interhuman communication; only if no such correlation can be found will FDG countenance the option of arbitrariness. In fact, languages can be shown to vary in the extent to which their formal properties reflect pragmatic or semantic categories or neither (cf. Hengeveld and Mackenzie 2008).

This position situates FDG halfway between radical formal and radical functionalist approaches. Radical functionalist positions tend to deny the existence of linguistic structure and see linguistic form as an ephemeral manifestation of the language user's attempt to achieve his/her communicative purposes. Radical formal positions contend that the utterances in an actual text or transcript of speech reflect (quite imperfectly, it is said) an underlying system that is governed by rules predicting the form taken by idealized linguistic units and limits linguistic study to the investigation of this covert system, totally independent of the uses to which it is put. FDG is a structural-functional theory (Butler 2003) in focusing on the correlation between function and structure, modeled as Formulation and Encoding respectively.

Two other structural-functional theories of language closely allied to FDG are Role and Reference Grammar (RRG; Van Valin and LaPolla 1997, Van Valin 2005, this volume) and Systemic Functional Linguistics (SFL; Halliday and Matthiessen 2004, Caffarel this volume); see Butler (2003) for detailed comparison. FDG appears to occupy a position intermediate between SFL, which stands closer to radical functionalism in taking the text to be the central object of linguistic investigation, and RRG, which stands closer to radical formalism in seeing itself as first and foremost a theory of syntax (Van Valin 2001*a*: 172). FDG has nothing to say about texts, but is very much concerned with the impact of textuality on the form of linguistic units; and FDG is not primarily interested in syntax, but does see morphosyntactic organization as one important aspect of linguistic encoding. With Simpler Syntax (Jackendoff and Culicover 2005, Culicover this volume) it shares the desire to give semantics its rightful place in linguistic theory and to integrate linguistics with cognitive, acquisitional, and language-biological work; it differs *inter alia* in giving equal weight to semantic and pragmatic factors.

FDG sees the language user as having knowledge of both functional and formal units and of the ways in which these units may be combined. This knowledge has a large degree of stability, such that it can be compared across languages, revealing universal trends in linguistic structure, as studied in language typology. This knowledge of units and their combination is instrumental in interpersonal communication and has arisen as a result of historical processes: formal and functional distinctions that have served human beings well through the ages have sedimented into the repertory now available to them. The forms that are at language users' disposal are variable across languages, but do not vary without limits. Rather, the limits are set by the range of communicative purposes displayed by all language users and by the cognitive constraints they are subject to.

This is the primary motivation behind the intimate relationship between FDG and linguistic typology. FDG is a theory that is capable of providing a framework for the enunciation and comparison of language universals (both absolute and statistical) and of offering a coherent model for the kind of language description that feeds into typological investigations. With its multi-layered structures of Formulation and Encoding, which define a space within which linguistic activity is constrained to operate, FDG permits more reliable comparisons of language systems. For example, FDG can readily accommodate the functionalist assumption that, *ceteris paribus*, the relative order of morphosyntactic elements will iconically reflect the scope relations holding among underlying pragmatic and semantic notions.

FDG offers a structured framework within which linguistic hypotheses can be enunciated and tested. At the same time, it provides a framework for the description of linguistic phenomena, and in this way can be involved in the entire cycle of research: from observation to prediction, to the testing of predictions through further observation, back to new predictions, and so on. FDG cannot in itself provide explanations, in the sense of rules of cause and effect. However, as we showed in section 15.2, it is linked to a Conceptual, a Contextual, and an Output Component, which themselves encompass all the linguistically relevant aspects of cognition, memory, and articulation; it is through these links that the extent of linguistic variation and its limitations can be made intelligible as reflecting general human mental and physical capacities.

15.4 Four levels of linguistic organization

15.4.1 Levels and layers

Each of the levels of representation distinguished within the Grammatical Component in Figure 15.2 is structured in its own way. What all the levels have in common is that they have a hierarchically ordered layered organization. In its maximal form the general structure of layers within levels is as follows:

(2) $(\pi\ v_1: [\text{head}\ (v_1)_\phi]: [\sigma\ (v_1)_\phi])$

Here v_1 represents the variable of the relevant layer, which is restricted by a (possibly complex) head that takes the variable as its argument, and may be further restricted by a modifier σ that takes the variable as its argument. The layer may be specified by an operator π and carry a function ϕ. Heads and modifiers represent lexical strategies, while operators and functions represent grammatical strategies. The difference between operators and functions is that the latter are relational, holding between the entire unit and other units at the same layer, while the former are not, applying only to the unit itself.

Not all relations between units are hierarchical. In those cases in which units together form a non-hierarchical (equipollent) configuration, they are enclosed between square brackets, as exemplified in (2), where the relationship between a head and its argument and a modifier and its argument is indicated by square brackets.

The levels differ as regards the nature of the distinctions that are relevant to each. Since the levels are purely linguistic in nature, only those distinctions are provided that are actually reflected in the grammar of the language involved. We will review the four different levels one by one, in the order that follows from the top-down organization of the model.

15.4.2 The Interpersonal Level

The Interpersonal Level captures all distinctions of Formulation that pertain to the interaction between Speaker and Addressee. These cover, at the higher layers, rhetorical notions of the overall structuring of discourse, to the extent that they are reflected in linguistic form, and at the lower layers, the pragmatic distinctions that reflect how Speakers mold their messages in view of their expectations of the Addressee's state of mind, again only to the extent that these are grammatically relevant. The hierarchical structure arises through the application of an appropriate set of frames from those available to the Speaker. The following shows the hierarchical relationships that apply at the Interpersonal Level:

(3)

$(\pi\ M_1$: [	Move
$(\pi\ A_1$: [	Discourse Act
$(\pi\ F_1\text{: ILL } (F_1)\text{: } \sum (F_1))$	Illocution
$(\pi\ P_1\text{: } \ldots\ (P_1)\text{: } \sum (P_1))_S$	Speaker
$(\pi\ P_2\text{: } \ldots\ (P_2)\text{: } \sum (P_2))_A$	Addressee
$(\pi\ C_1$: [	Communicated Content
$(\pi\ T_1\text{: } [\ldots]\ (T_1)\text{: } \sum (T_1))_\Phi$	Ascriptive Subact
$(\pi\ R_1\text{: } [\ldots]\ (R_1)\text{: } \sum (R_1))_\Phi$	Referential Subact
$]\ (C_1)\text{: } \sum (C_1))_\Phi$	Communicated Content
$]\ (A_1)\text{: } \sum (A_1))_\Phi$	Discourse Act
$]\ (M_1)\text{: } \sum (M_1))$	Move

We will now say something about each of the layers in turn.

The Move (M_1) is the largest unit of interaction relevant to grammatical analysis. It may be defined as an autonomous contribution to the ongoing interaction: it either calls for a reaction or is itself a reaction. The complexity of a Move may vary enormously, from silence through to a lengthy stretch of discourse. Where linguistic material is present, the Move will always take the form of one or more Discourse Acts. Its general frame is as follows:

(4) $(\pi\ M_1\text{: } [(A_1)\ldots(A_{1+N})]\ (M_1)\text{: } \sum (M_1))$, where $N \geq 0$

The relationship between the Discourse Acts may be one of equipollence or of dependence. Prominent relationships of dependence, indicated as a rhetorical function on the dependent Discourse Act, are Motivation, Concession, Orientation, and Correction. In the following Move:

(5) Watch out, because there will be trick questions in the exam.

the second (intonationally distinct) Discourse Act with a Declarative Illocution serves to indicate the Speaker's motivation for uttering an Imperative Illocution in the first Discourse Act.

The representation of a Discourse Act will show only those components that have actually been deployed by the Speaker, minimally the Illocution (F_1) and the Speaker $(P_1)_S$ him/herself. Three kinds of Discourse Acts are distinguished:

(6) Expressives, which give direct expression to the Speaker's feelings
e.g., *Ouch!* (A_I: [(F_I: /aʊtʃ/$_{Int}$ (F_I)) $(P_I)_S$] (A_I))
Interactives, which consist of invariable, often ritualized lexical material
e.g., *Congratulations!* (A_I: [(F_I: /kəŋgrætju:ˈleɪʃnz/(F_I)) $(P_I)_S$ $(P_J)_A$] (A_I))
Contentives, which involve a Communicated Content and either a lexical or abstract Illocution (F_1)
e.g., *I promise to be there tomorrow* (A_I: [(F_I: /ˈprɒmɪs/$_V$(F_I)) $(P_I)_S$ $(P_J)_A$ (C_I)] (A_I))
I'll be there tomorrow (A_I: [(F_I: DECL(F_I))$(P_I)_S$$(P_J)_A$ (C_I)] (A_I))

Discourse Acts can be modified lexically, for example by an expression indicating the style of the Act (*briefly*). They may also be subject to operators, such as those for emphasis, irony, and mitigation.

The head of the Illocution may be either lexical or abstract, as already illustrated in (6). This also applies to Vocative Illocutions, for example allowing an analysis of the epistolary salutation *Dear John* as:

(7) (A_I: [(F_I: /dɪə/ (F_I)) $(P_I)_S$ (P_J: /dʒɒn/$(P_J))_A$] (A_I))

Typical modifiers of Illocutions are illocutionary adverbs such as *honestly*, as in:[2]

(8) *Honestly*, I don't like you.

(9) (M_I: [(A_I: [(F_I: DECL (F_I): –honestly– (F_I)) $(P_I)_S$ $(P_J)_A$ (C_I: –I don't like you– (C_I))] (A_I))] (M_I))

The two Participants in an interaction, (P_1) and (P_2), alternate as Speaker and Addressee; these roles are therefore indicated as functions. The head may be abstract (and left unexpressed) or may be lexical, as in (10) and (11):

(10) *The company* hereby undertakes to replace any can of Doggo-Meat that fails to please, with no questions asked. (Levinson 1983: 260)

[2] Note that in cases in which not all details are necessary for the analysis of the phenomenon at hand, we use the symbol "–" to indicate the beginning and the end of a fragment that is not further analyzed in detail.

(11) Japanese (Hinds 1986: 257)

Iroiro-to	*suwan*	*san*	ni	shitsumon	shimasu.
various	Swan	Ms	REC	question	do

'I'd like to ask you a variety of questions, Ms Swan.'

The Communicated Content (C_1) contains the totality of what the Speaker wishes to evoke in his/her communication with the Addressee. Communicated Contents have their distinctive operators and modifiers. One operator that has received attention in FDG is the reportative, which must be distinguished from the evidential operators of the Representational Level. Each (C_1) contains one or more Subacts, so called because they are hierarchically subordinate to Discourse Acts. Subacts bear pragmatic functions, and the frames for Communicated Contents ("content frames") are shown as configurations of these pragmatic functions, e.g., as thetic, categorical, etc.

FDG recognizes three pragmatic functions, which are assigned only when relevant (i.e., where they have an impact on linguistic form). The Focus function signals the Speaker's strategic selection of new information, either to fill a gap in the Addressee's information or to correct that information. The segment of (C_1) not assigned the Focus function constitutes the Background. The Topic function is assigned to a Subact which has a special function within the Discourse Act, that of signaling how the Communicated Content relates to the gradually constructed record in the Contextual Component. The segment not assigned the Topic function constitutes the Comment. It is typically the Focus and/or Topic that are encoded in languages; formal expression of Background and Comment is rare. Languages may lack the Topic function, or permit multiple Topic and/or Focus. A third pragmatic function is Contrast (as opposed to Overlap), which signals the Speaker's desire to bring out the differences between two or more Communicated Contents or between a Communicated Content and contextually available information. The three functions may in principle be combined with each other, and indeed we find Focus/Contrast combinations in English cleft constructions, Topic/Contrast in Korean NPs marked by *-nɯn* (Lee 1999) and Focus/Topic in such presentative constructions as French (12):

(12)

Il	est	arrivé	trois	trains.
it	AUX.PRS.3.SG	arrive.PTCP.SG.M	three	trains

"There arrived three trains."

There are two types of Subact: an Ascriptive Subact (T_1) is an attempt by the Speaker to evoke a property, while a Referential Subact is an attempt by the Speaker to evoke a referent. In certain languages, e.g., Samoan (Mosel and Hovdhaugen 1992) and Tagalog (Himmelmann 2008), the (T) or (R) status of Subacts is marked explicitly. The head of a (T_1) is in principle empty (the Property being indicated at the Representational Level), but it may be modified by items such as *allegedly*, *fortunately*, *really* and/or may be subject to an approximative operator, expressed in English as *sort-of*, typically /sɔ:də/. The head of an (R_1) is typically itself an Ascriptive Subact (as in *the hat*), but may be a Proper name (*Mary*) or an abstract head (realized as a pronoun or affix). Among the modifiers of Referential Subacts are forms such as *poor* in (*No-one cares about*) *poor me*; and the principal operators are those for specificity (±s) and identifiability (±id). A special case is the combination {+id, −s}, which may be associated with Evans's (2003) notion of the ignorative, where the referent is assumed identifiable for the Addressee but not for the Speaker.

15.4.3 The Representational Level

The Representational Level deals with the semantic aspects of a linguistic unit. Whereas the Interpersonal Level takes care of evocation, the Representational Level is responsible for designation. The use of the term "semantics" is thus restricted to the ways in which language relates to the possible worlds it describes. The layers relevant at the Representational Level are defined in terms of the semantic categories they designate. Semantic categories are the language-specific linguistically relevant manifestations of ontological categories. They are hierarchically organized as indicated in (13):

(13)

$(\pi\ p_1:$	Propositional Content
$\quad(\pi\ ep_1:$	Episode
$\quad\quad(\pi\ e_1:$	State-of-Affairs
$\quad\quad\quad[(\pi\ f_1: [$	Configurational Property
$\quad\quad\quad\quad\quad(\pi\ f_1: \blacklozenge\ (f_1):\ [\sigma\ (f_1)_\Phi])$	Lexical Property
$\quad\quad\quad\quad\quad(\pi\ x_1: \blacklozenge\ (x_1):\ [\sigma\ (x_1)_\Phi])_\Phi$	Individual
$\quad\quad\quad\quad\quad\ldots$	
$\quad\quad\quad]\ (f_1): [\sigma\ (f_1)_\Phi])$	Configurational Property
$\quad\quad(e_1)_\Phi]:\ [\sigma\ (e_1)_\Phi])$	State-of-Affairs
$\quad(ep_1):\ [\sigma\ (ep_1)_\Phi])$	Episode
$(p_1):\ [\sigma\ (p_1)_\Phi])$	Propositional Content

Propositional Contents (p), the highest units at the Representational Level considered here, are mental constructs, such as pieces of knowledge, beliefs, and hopes. Propositional contents may be factual, as when they are pieces of knowledge or reasonable belief about the actual world, or non-factual, as when they are hopes or wishes with respect to an imaginary world. Given their nature, Propositional Contents are characterized by the fact that they may be qualified in terms of propositional attitudes (certainty, doubt, disbelief) and/or in terms of their source or origin (shared common knowledge, sensory evidence, inference). Propositional Contents

(p) are not identical to Communicated Contents (C), which were discussed in the previous section. Communicated Contents constitute the message contents of Discourse Acts, and are not necessarily propositional in nature. Thus, though the Communicated Content of an act may correspond to a Propositional Content, it is not identical to it. A major difference between Communicated Contents and Propositional Contents is that the former are Speaker-bound, whereas the latter are not, at least not necessarily. This means that Propositional Contents can be attributed without problems to persons other than the Speaker:

(14) Jenny believed that/hoped that/went home because *maybe* her mother would visit her.

In all these examples the embedded Propositional Content is attributed to the Individual *Jenny* introduced in the main clause. The propositional nature of the parts in italics in (14) shows up in the fact that it may contain elements expressing a propositional attitude, such as *maybe*.

Propositional Contents contain Episodes (ep), which are sets of States-of-Affairs that are thematically coherent, in the sense that they show unity or continuity of Time (t), Location (l), and Individuals (x). In various languages the semantic category of Episodes is very manifestly present in the grammatical system, for instance in those that exhibit Tail-Head linkage. But we also need it for English sentences like the following one, adapted from Givón (1995; see also Wanders in prep.):

(15) Coming out, stopping to check the mailbox, taking a look at the driveway, and pausing to adjust his hat, he walked to his car.

Here a series of concatenated non-finite narrative verb forms, together with a final finite verb form, together describe an Episode within a larger story. The example at the same time shows an important aspect of Episodes: they are located in absolute time, while States-of-Affairs are located in relative time. Thus, while all the clauses in (15) represent States-of-Affairs, absolute location in time occurs only once for the series as a whole.

States-of-Affairs (e) include events and states and are characterized by the fact that they can be located in time and can be evaluated in terms of their reality status. States-of-Affairs can thus be said to "(not) occur", "(not) happen", or "(not) be the case" at some point or interval in time. The following example shows once more that absolute time, a feature of Episodes, may combine very well with relative time, a feature of States-of-Affairs:

(16) Yesterday Sheila went out before having dinner.

The absolute setting provided by the adverb *yesterday* holds for the two States-of-Affairs contained within (16) as they form part of the same Episode. The adposition *before* specifies the relative temporal relation between the two.

Some languages mark this distinction in their grammatical systems. The following example is from Swahili (Ashton 1944: 133). In this case the first verb form provides the absolute temporal setting, while subsequent narrative verb forms indicate relative chronological subsequence:

(17)	Ni-li-kwenda	soko-ni,	ni-ka-nunua	ndizi	sita,
	1.SG-PST-go	market-LOC	1.SG-SUBS-buy	banana	six,
	ni-ka-la	tatu,	ni-ka-mpa	mwenz-angu	tatu.
	1.SG-SUBS-eat	three	1.SG-SUBS-give	companion-1.SG.POSS	three

'I went to the market, and bought six bananas; I ate three and three I gave to my companion.'

After indicating that the first State-of-Affairs in the series occurred in the past by using the prefix *li-*, the remaining States-of-Affairs within the Episode can be marked as having taken place subsequent to the last-mentioned State-of-Affairs by means of the prefix *ka-*.

A State-of-Affairs is characterized by a Configurational Property (f), which is compositional in nature and contains a combination of semantic units that are not in a hierarchical relationship with respect to each other. Configurational Properties constitute the inventory of predication frames relevant to a language. Languages may differ markedly in the nature and number of predication frames that are allowed with respect to both their quantitative and their qualitative valency. As for quantitative valency, there may for instance be restrictions on the maximum valency that a language allows in combination with a single predicate. In many serializing languages the maximum valency of a verb is 2, and serialization is required to expand that valency indirectly, as in the following example from Mandarin Chinese (Li and Thompson 1981: 366):

(18)	Wǒ	gěi	nǐ	dào	chá.
	I	give	you	pour	tea

'I'll pour you some tea.'
'I pour tea give you.'

Qualitatively speaking, languages may, for instance, differ as regards the division of labor between semantic functions. Thus in Tariana no distinction is made between the formal encoding of ablative, essive, and allative (Aikhenvald 2003: 148):

(19)	Na-pidana	uni-*se*.
	3.PL.go-REM.PST.REP	water-LOC

'They went into water.'

(20)	Nawiki	pa:-putʃita-*se*	nehpani-pidana.
	people	one-CL-LOC	3.PL.work-REM.PST.REP

'People were working on a clearing.'

(21) Hĩ wyaka-*se* ka-nu-kaɾu dhuma-naka
DEM.ANIM far-LOC REL-come-PST.REL.F 3.SG.F.hear-PRS.VIS
waku-nuku.
1PL.speech-TOP
'She who came from far away understands our speech.'

Configurational Properties are built up using semantic categories that are in a non-hierarchical relationship with one another. These semantic categories may be of various types and include Individuals (x), i.e., concrete objects that can be located in space, and Lexical Properties (f), which have no independent existence and can only be evaluated in terms of their applicability to other types of entity. Further semantic categories may be relevant to the grammar of an individual language and enter into the constitution of a Configurational Property, such as Location (l), Time (t), Manner (m), Reason (r), and Quantity (q). In all cases, only those semantic categories are postulated for a language that trigger formal processes within the grammar of that language. By way of example, consider the English nominalization strategies exemplified in Table 15.1. English has distinct nominalization processes that create nouns designating Propositional Contents, States-of-Affairs, Properties, Individuals, and Locations. The postulation of these semantic categories within the grammar of English is thus warranted on formal grounds.

15.4.4 The Morphosyntactic Level

The Morphosyntactic Level deals with the structural aspects of a linguistic unit. Together with the Phonological Level, it takes care of the encoding of interpersonal and representational distinctions. In view of this function, much of what happens at the Morphosyntactic Level is functionally motivated: ordering principles are motivated by iconicity, domain integrity, and the preservation of scope relations. At the same time, morphosyntax has its own principles of organization, as for instance in the arbitrary imposition of a basic constituent order pattern, which in itself cannot be argued to be functionally motivated. FDG does not make a distinction

Table 15.1. Derived nominal expression of basic semantic categories

Entity type	Examples
p	*hope-Ø, wish-Ø, belief-Ø*
e	*explora-tion, deci-sion, deple-tion*
f	*mean-ness, kind-ness, false-ness*
x	*writ-er, employ-er, sing-er*
l	*bak-ery, brew-ery, eat-ery*

between a syntactic and a morphological level of analysis, since the principles used in the formation of words are the same as those used in the formation of phrases and clauses.

The layers relevant at the Morphosyntactic Level are listed in (22):

(22)	$(Le_1$:	Linguistic Expression
	$(Cl_1$:	Clause
	$(Xp_1$:	Phrase
	$(Xw_1$:	Word
	(Xs_1)	Stem
	(Aff_1)	Affix
	$(Xw_1))$	Word
	$(Xp_1))$	Phrase
	$(Cl_1))$	Clause
	$(Le_1))$	Linguistic Expression

A Linguistic Expression is any set of at least one morphosyntactic unit; where there is more than one unit within a Linguistic Expression, these will demonstrably belong together in their morphosyntactic properties. The units combining into a Linguistic Expression may be Clauses, Phrases, or Words. The following German example illustrates a combination of Phrases:

(23) Je kürzer desto besser.
CORR short.COMPV CORR good.COMPV
'The shorter the better.'

Here we have two mutually dependent Adjective Phrases linked by the correlative pair *je ... desto*, thus illustrating a Linguistic Expression which does not contain a Clause:

(24) $(Le_i$: $[(Ap_i$: $[(Gw_i$: je $(Gw_i))$ $(Aw_i$: kurz-Compv $(Aw_i))]$ $(Ap_i))$
$(Ap_j$: $[(Gw_j$: desto $(Gw_j))(Aw_j$: gut-Compv $(Aw_j))]$ $(Ap_j))]$ $(Le_i))$

By introducing Linguistic Expressions as the highest category in its morphosyntax, FDG creates a possibility of dealing straightforwardly with holophrases and non-sentential utterances.

A simple Clause is a grouping of one or more Phrases and possibly (grammatical) Words and is characterized, to a greater or lesser extent, by a template for the ordering of those Phrases and, also to a greater or lesser extent, by morphological expressions of connectedness (notably government and agreement); in addition, the Clause may operate as a domain for several morphosyntactic processes. While for each language analyzed, the identification of Clauses will be dependent upon language-specific criteria, we believe that it is justified to posit the Clause as a universal category of morphosyntactic structure.

A Phrase (Xp) is headed by a lexical item that is passed on from the Interpersonal Level or the Representational Level. There is no necessary one-to-one correspondence between the lexeme classes recognized in a language and the Phrase types and corresponding Word classes recognized within that same language. A language with a highly flexible lexeme class may have a variety of Phrase types. Consider the following example from Mundari (Evans and Osada 2005: 354–5):

(25) Buru=ko bai-ke-d-a.
mountain=3PL make-COMPL-TR-PRED
'They made the mountain.'

(26) Saan=ko buru-ke-d-a.
firewood=3PL mountain-COMPL-TR-PRED
'They heaped up the firewood.'

The lexeme *buru* can be used as the head within a Referential Subact (25) and as the head within an Ascriptive Subact (26), and can thus be characterized as a flexible lexeme. Yet the morphosyntax of Mundari makes a clear distinction between the Phrase expressing the Ascriptive Subact and the one expressing the Referential Subact, traditionally called "Verb Phrase" and "Noun Phrase":

(27) $(\mathrm{Np_i}: (\mathrm{Nw_i}: \mathrm{buru_{Cont}}\ (\mathrm{Nw_i}))\ (\mathrm{Np_i}))$

(28) $(\mathrm{Vp_i}: (\mathrm{Vw_i}: \mathrm{buru_{Cont}}\ (\mathrm{Vw_i}))\ (\mathrm{Vp_i}))$

The Nominal and Verbal Word templates will then be different as regards their possibilities for suffixation.

The Word itself (Xw), especially in incorporating languages, can be highly complex. Apart from the fact that it may consist of Stems (Xs) and Affixes (Aff), in some languages it may, just like any other layer of morphosyntactic analysis, embed higher layers such as phrases and clauses, obeying full recursivity. Consider the following example from Chukchee (Skorik 1961: 103, discussed in Mattissen 2006: 290):

(29) Tə-[tor-taŋ-pəlwəntə-pojgə]-pela-rkən.
1.SG.ABS-new-good-metal-spear-leave-PRES.1.SG>3.SG
'I am leaving a good, new, metal spear.'

In this example a Noun Phrase, including its modifiers, is incorporated as a whole within the Verbal Word and is crossreferenced on the verbal word itself. Together these facts point to the phrasal status of the incorporated noun and its modifiers, as represented in (30):

(30) $(\mathrm{Vw_i}: [(\mathrm{Aff_i}: \text{tə}\ (\mathrm{Aff_i}))(\mathrm{Np_i}: \text{tortaŋpəlwəntəpojgə}\ \ (\mathrm{Np_i}))(\mathrm{Vs_i}: \ \text{pela}\ (\mathrm{Vs_i}))$
$(\mathrm{Aff_j}: \ \textsc{pres}\ (\mathrm{Aff_j}))]\ (\mathrm{Vw_i}))$

Each internally complex layer of morphosyntactic analysis is built up in a number of steps. The linear order of elements is considered from two different perspectives.

As we argued earlier, the Interpersonal and Representational Levels are partially organized hierarchically and partially configurationally. Linear ordering starts out with the hierarchically higher elements and works down to the lower ones, in line with the top-down organization of the model as a whole. This initial step implements the fact that hierarchical scope relations are reflected in linear order. Interpersonal and representational units that are in a configurational relationship cannot be ordered in this way. In order to determine how these should be placed with respect to each other the alignment system of the language now comes into play. Alignment may be based on interpersonal, representational, or morphosyntactic triggers, or a combination of these. All linear ordering is done dynamically, by making use of a number of absolute positions (maximally Initial, Second, Medial, and Final). Once these have been filled, positions relative to them become available.

Obligatory positions in templates of any layer for which no material is available from the Interpersonal and Representational Levels are filled with syntactic or morphological dummies in a process that is called coercion. Thus, in many languages the insertion of a non-verbal constituent in the predicate slot will trigger the insertion of a copula. In others, the insertion of a basically transitive lexeme in an intransitive predication frame will trigger a detransitivizing affix. Once all the positions in a template have been filled, a number of copying operations may be necessary in order to account for the expression of agreement, sequence of tenses, and the like.

15.4.5 The Phonological Level

The Phonological Level is responsible for every aspect of Encoding not covered by the Morphosyntactic Level. It receives input—some of it already in phonemic form—from all three other levels and provides input to the Output Component. Whereas the latter deals with such “analog” matters as formant frequency, intensity, duration and spectral characteristics, the Phonological Level—being within the grammar—is “digital”, containing representations in phonemes that are ultimately based in binary phonological oppositions. In other words, the Phonological Level does not show the “melody” of the Intonational Phrase but provides a number of indications at each layer which the Output Component converts into a smoothly flowing result.

The primitives with which the Phonological Level operates include (i) the prosodic patterns that apply at each layer of analysis; (ii) an inventory of segmental sequences (the “grammatical lexicon”) expressing particular configurations of morphemes or placeholders introduced at other levels; and (iii) a set of tertiary operators which will have their ultimate effect in the Output Component.

Just like the other levels, phonological representations are hierarchical in nature (as in the tradition of Prosodic Phonology initiated by Nespor and Vogel 1986). Here, too, FDG makes the assumption that not all layers are active in every Utterance or indeed are relevant to every language system. And as at the Morphosyntactic Level, FDG does not exclude the possibility of recursion at certain layers. The maximum layering of the Phonological Level is as follows:

(31)	$(\pi\mathrm{U}_1$: [	Utterance
	$(\pi\mathrm{IP}_1$: [	Intonational Phrase
	$(\pi\mathrm{PP}_1$: [	Phonological Phrase
	$(\pi\mathrm{PW}_1$: [	Phonological Word
	$(\pi\mathrm{F}_1$: [	Foot
	$(\pi\mathrm{S}_1)$	Syllable
	] $(\mathrm{F}_1))$	Foot
	] $(\mathrm{PW}_1))$	Phonological Word
	] $(\mathrm{PP}_1))$	Phonological Phrase
	] (IP_1)	Intonational Phrase
	] $(\mathrm{U}_1))$	Utterance

We will now say something about each of the layers in turn.

The Utterance (U_1) is the largest stretch of speech covered by the Phonological Level. A Speaker will tend to use more substantial pauses to separate Utterances than Intonational Phrases; these longer pauses will also never be interpreted by the Addressee as hesitations (Hayes 1989: 219). An Utterance may in addition display pitch distinctions called paratones which help to mark it off as a self-contained group of Intonational Phrases (Brown and Yule 1983: 101); FDG represents these as operators on the (U)-variable. The Output Component may react to an Utterance boundary by introducing such phenomena as “final F0 lowering, segmental lengthening, creaky voice, amplitude lowering, long pauses, stylized ‘finality’ contours, etc.” (Venditti 2005: 191).

The Intonational Phrase is characterized by a nucleus, i.e., a pitch movement localized on one or more Syllables which is essential to the interpretation of the Intonational Phrase as a whole; FDG represents this global pitch movement as an operator—(f)alling, (r)ising—on the IP variable, cf. (32b) below. One Intonational Phrase is typically separated from another by a pause (shorter than that between Utterances); in the Output Component there may be additional rhythmic or durational indications. The gradual integration of Discourse Acts within a Move may be reflected in the loss of (IP) boundaries within the Utterance. In (32), the amalgamation of an Orientation and a Nuclear Discourse Act, as in (32b), induces an analysis with a single Intonational Phrase, as in (32c) (a French example inspired by Di Cristo 1998: 211):

(32) a. Mon voisin il est toujours malade.
1SG.POSS neighbor 3SG.M be.PRS.3SG always ill
'My neighbor, he's always ill'; or: 'My neighbor is always ill.'
b. ((U$_i$: [(rIP$_i$: /mɔ̃vwazɛ̃/(IP$_i$)) (fIP$_j$: /ilɛtuʒuʀmalad/(IP$_j$))] (U$_i$))
c. ((U$_i$: (fIP$_i$: /mɔ̃vwazɛ̃ilɛtuʒuʀmalad/(IP$_i$)) (U$_i$))

The Phonological Phrase in stress languages contains one Syllable that is more strongly stressed than the others; this Nuclear Syllable is typically also the primary location for the global fall or rise within the Intonational Phrase. In tone languages, in which pitch movement is used for lexical distinctions, Phonological Phrases have a different *raison d'être*, namely as the domain of tone sandhi. In the stress language English, both DECL and IMP Illocutions are characterized by a globally falling pitch at the layer of the Intonational Phrase (fIP$_i$). However, the fall on the Nuclear Syllable tends to be much more marked with the IMP Illocution; this is indicated by assigning an additional falling tertiary operator to the (PP$_1$) containing the Nuclear Syllable. The Output Component interprets such a double indication of fall as entailing a larger pitch movement downward. Pitch height (as opposed to movement) within the Phonological Phrase—(h)igh, (m)id, (l)ow—is in very many languages associated with the expression of pragmatic functions; see 15.5.2.4 for an example.

The Phonological Word (PW$_1$), for those languages in which such a category needs to be recognized, is a slice of phonological structure which displays at least one criterial characteristic, which may relate to the number of segments, to prosodic features, or to the domain of phonological rules. Its complex relation to the Morphosyntactic Word will be treated in section 15.5.2.7. One of the principal tasks of the Phonological Level is to convert all placeholders from other levels into phonological form and to integrate them into a Phonological Word. To achieve this, the Phonological Level has a store of primitives at its disposal which provide phonemic material with which to replace the placeholders in the input. This store of primitives constitutes the "grammatical lexicon" of the language under analysis. An example is the English comparative, where the form depends on the phonological characteristics of the Adjective (number of syllables and stress placement): the lexical item *more* therefore appears as a placeholder at the Representational and Morphosyntactic Levels, the final choice between the Phonological Word /mɔː/ and the Syllable /-ə/ being determined at the Phonological Level.

Phonological Words are divided into Syllables, which in stress languages (i.e., those with stressed and unstressed Syllables) group into Feet. Stress is indicated by the operator "s" on the Syllable variable. Non-accentual tone (e.g., in Thai), tone accent (e.g., in Swedish), and accentual tone (e.g., in Japanese) similarly involve operators—i.e., the position π—on (π s$_1$).

15.5 Interplay between the components and levels

15.5.1 Relations between components

15.5.1.1 *Introduction*

As was made clear in section 15.2, the Grammatical Component described in section 15.4 is part of a wider theory of verbal interaction. The architecture proposed for this theory in FDG work is strongly inspired by the extensive research into the processes of speech production detailed in Levelt (1989). His model distinguishes three fundamental modules: the Conceptualizer, the Formulator, and the Articulator. Very roughly, these correspond to our Conceptual Component, Grammatical Component, and Output Component respectively; to these FDG has added a Contextual Component. We will discuss the interactions between these components one by one.

15.5.1.2 *Interplay between the Grammatical and Conceptual Components*

The Conceptual Component is the driving force behind the workings of the Grammatical Component. It is here that is represented the ideational and interactive material that motivates Discourse Acts and the Moves in which they occur. The Conceptual Component does not include every aspect of cognition but only those that reflect the immediate communicative intention. For example, a Speaker may conceive the desire to impart some bad news to the Addressee and concurrently to show sympathy. In English, the Conceptual Component can trigger the operation of Formulation to structure this as a Move with two Discourse Acts, one with a Declarative Illocution, the other containing an appropriate Interactive formula at the Interpersonal Level:

(33) $(M_I: [(A_I: [(F_I: \text{DECL } (F_I)) (P_I)_S (P_J)_A (C_I: [(T_I) (R_I)] (C_I))] (A_I))$
$(A_J: [(F_J: /\text{aɪmə'freɪd}/(F_J)) (P_I)_S (P_J)_A] (A_J))](M_I))$
"John's ill, I'm afraid."

In (32), the distinct ideational and affective-interactional elements (cf. Butler 2008) are reflected in separate Discourse Acts. An alternative is for the Conceptual Component to trigger a single Move, as in *I'm afraid John's ill*, a Move with only one Discourse Act which is simultaneously a Declarative and an expression of sympathy. Here *I'm afraid* will appear as a Modifier of the Discourse Act:

(34) $(M_I: [(A_I: [(F_I: \text{DECL } (F_I)) (P_I)_S (P_J)_A (C_I: [(T_I) (R_I)] (C_I))] (A_I):$
$(F_J: /\text{aɪmə'freɪd}/(F_J)) (A_I))] (M_I))$
"I'm afraid John's ill."

Although the Conceptual Component is ancillary to the Grammatical Component, it does not cover the same as Slobin's (1996) notion of "thinking for speaking". Whereas that notion is language-specific and involves "picking those characteristics of objects and events that (i) fit some conceptualization of the event, and (ii) are readily encodable in the language" (Slobin 1996: 76), the Conceptual Component is pre-linguistic. The kinds of examples Slobin gives, e.g., the witnessed/non-witnessed opposition in Turkish or the perfective/imperfective distinction in Spanish, are in FDG grammatical choices that are determined through the operation of Formulation.

15.5.1.3 *Interplay between the Grammatical and Contextual Components*

Just as the Conceptual Component is limited in its compass, so the Contextual Component, too, does not aim to represent the entire ongoing discourse but rather to house just those aspects of the context that impinge upon the workings of the Grammatical Component. Thus it contains all the information from the grammar that is relevant to the form taken by subsequent utterances; and it stores longer-term information about the current interaction that is relevant to Formulation and Encoding in the language being used. As examples of the latter, consider the fact that in languages like Spanish knowledge of the sex of the speech act participants and the social relation between them is essential for interaction. In (35), the choice of the forms *pálida* (rather than *pálido* "pale-M.SG") and *estás* (rather than *está* "COP-IND.PRS.2.SG.POL") reflects specifications in the Contextual Component:

(35) ¡Qué pálid-a est-ás!
what pale-F.SG COP-IND.PRS.2.SG.FAM
'How pale you look!'

For an account of the grammatical properties of the corresponding utterance in English, as in the translation of (35), no such specification is required.

As examples of grammatical phenomena that presuppose the first-mentioned function of the Contextual Component, consider reflexives, anaphora, and instances of narrative chaining. In languages with logophoric pronouns, for example, the Contextual Component will have to keep track of the status of (typically human) entities as belonging to a particular embedded discourse domain or not. Similarly, according as a language permits reflexive pronouns to apply across larger or smaller stretches of discourse, the Contextual Component will be adjusted to make particular possible antecedents available. The Contextual Component keeps track not only of the results of Formulation but also from those of Encoding, since anaphoric reference is possible not only to pragmatic and semantic constructs but also to sections of the actual morphosyntactic structure of linguistic expressions and the phonological structure of utterances.

15.5.1.4 *Interplay between the Grammatical and Output Components*

The function of the Output Component in speech may be seen as translating the digital (i.e., opposition-based) information in the grammar into analog (i.e., continuously variable) form. An Utterance boundary at the PL will accordingly yield a pause of so many milliseconds in the Output Component; or a Syllable with a "falling" operator will bring about a decline in the fundamental frequency of the corresponding stretch of the output. The Output Component is also the location for long-term settings, such as the tempo at which an individual's speech is carried out: *allegro* forms attributable to fast speech are among the phenomena treated here.

As an example of the effect of the Output Component, consider degemination (cf. Booij 1999*b*: 68–9, 151). In Dutch (but not for example in English), there is a requirement that two identical adjacent consonants (or consonant clusters like /st/) be reduced to one. This can apply within lexical compounds, such as *kunststuk* "objet d'art" /ˈkœnstœk/; cf. *kunst* "art" /kœnst/ and *stuk* "piece" /stœk/; the lexical entry already shows the effect of degemination. It can also apply to the result of morphosyntactic processes, for example with the sequence /zɪt/ + /t/ "sit + 3.SG.PRES" being realized as /zɪt/ at the Phonological Level (*/zɪt:/). However, chance adjacencies can also happen inside Intonational Phrases, as for example within the Phonological Phrase in the analysis of (36):

(36) zit-Ø te werk-en.
sit-3SG.PRES CNJ work-INF
'is working.'
(PP$_i$: (PW$_i$: ([(SS$_i$: /zɪt/(S$_i$))(S$_j$: /tə (S$_j$))](PW$_i$))(PW$_j$: ([(SS$_k$: /ʋɛʁ/(S$_k$))(S$_l$: /kən/(S$_l$))](PW$_i$))](PP$_i$))

Here it is the Output Component that imposes the degemination, yielding an output transcribable as [zɪtəʋɛʁkə].

15.5.2 Relationships between Levels of the Grammatical Component

15.5.2.1 *Introduction*

Even a glimpse at the layered hierarchies of the four grammatical levels suggests that there is a high degree of correspondence among them, and there are indeed default correlations between, for example, Discourse Act, State of Affairs, Clause and Intonational Phrase, or between Subact, Property/Individual, Phrase and Phonological Phrase. These correlations are far from perfect, however, and differ across languages as well. In the following subsections, we will briefly consider the relations across the various layers.

15.5.2.2 *Relationship between the Interpersonal and Representational Levels*

Every linguistic item is analyzed at the Interpersonal Level: as we saw in section 15.4.2, even Expressives, despite not being other-directed, involve a Speaker and an Illocution. Only if the Interpersonal Level contains a Communicated Content does the Representational Level also come into play. In these cases, the two Levels share responsibility for Formulation. In an example such as (37), *Mary*, its Modifier *poor*, and *really* originate at the Interpersonal Level, while the Representational Level introduces *like* and *seem* into appropriate frames and embeds the units encoded as the extraposed Subject of *seem* and the infinitive complement of *like*:

(37) It seems poor Mary really likes to suffer.

Although Ascriptive Subacts (T_1) typically correspond to Properties (as in Turkish (38)), the (T_1) in English (39) is an Individual at the RL:

(38) Erkek öğretmen-Ø-Ø.
man teacher-PRES-3SG
'The man is a teacher.'
IL: C_I T_I R_I
RL: $(p_i: -(e_i: [\ (f_i: \text{öğretmen}_N\ (f_i))\ (x_i: (f_j: \text{erkek}_N\ (f_j))\ (x_i))_{\emptyset}]\ (e_i))-(p_i))$

(39) The man is a teacher.
IL: C_I T_I R_I
RL: $(p_i: -(e_i: [(x_i: (f_i: \text{teacher}_N\ (f_i))\ (x_i))\ (x_j: (f_j: \text{man}_N\ (f_j))\ (x_j))]\ (e_i))-(p_i))$

15.5.2.3 *Relationship between the Interpersonal and Morphosyntactic Levels*

The distinctions drawn at the Interpersonal Level are encoded at either the Morphosyntactic or the Phonological Level. Let us consider the encoding of the pragmatic function Focus assigned to Subacts. Although Focus is in many languages associated with prosodic effects at the Phonological Level, it can also be encoded morphosyntactically. For example, in Tariana (Aikhenvald 2003: 139), the suffix *-nhe/-ne* is applied to Subjects in Focus (to simplify a little); other languages have markers for Focus alone, cf. Wambon *-nde* (de Vries 1985: 172). Another way of marking Focus morphosyntactically is by assigning a special syntactic position to the Focus element: in Aghem (Watters 1979: 144) it is the immediately post-verbal position, in Hungarian (Kenesei et al. 1998: 166) the immediately pre-verbal position. Finally, many languages indicate Focus with a special Focus construction: where this takes the form of a cleft construction, this strategy involves the Representational Level as well, dividing the content into two segments, with one part (the one corresponding to the Focus element) being predicated of the other.

15.5.2.4 *Relationship between the Interpersonal and Phonological Levels*

Although they are maximally separated in the model, the relationship between the Interpersonal and Phonological Levels is very close. As was mentioned in 15.4.2, Focus is in many languages—iconically—associated with phonological prominence, as are the other pragmatic functions. Illocutionary distinctions also tend to be expressed phonologically, especially if there is no morphosyntactic indication: in Portuguese, for example, the distinction between Declarative and Interrogative is signaled only through an opposition between a falling and rising operator, respectively, on the Intonational Phrase: these have their effect on its final Phonological Phrase.

In English, the syntax of the Clause is usually geared to ensuring Clause-final placement for the element associated with Focus assignment; the default effect on the Phonological Level is for the final Phonological Phrase to indicate both the Illocution and the placement of the Focus, as in (40):

(40) a. I saw [a heron]$_{Foc}$.
b. (fIP$_i$: [(PP$_i$: /aɪˈsɔː/ (PP$_i$)) (PP$_j$: /əˈherən/ (PP$_j$))] (IP$_i$))

In an example such as (41), the entire Communicated Content is in Focus:

(41) a. [[The train] arrived]$_{Foc}$.
b. (fIP$_i$: [(PP$_i$: /ðəˈtreɪn/ (PP$_i$)) (lPP$_j$: /əˈraɪvd/ (PP$_j$))] (IP$_i$))

The f-operator on (IP$_i$) would normally induce a falling intonation on the Syllable /raɪvd/; however, this is rendered impossible by the presence of the l(ow)-operator on (PP$_j$). The Output Component will therefore apply a fall to the preceding (PP), and the pitch will continue low.

15.5.2.5 *Relationship between the Representational and Morphosyntactic Levels*

The relationship between the Representational and Morphosyntactic Levels is guided by the principle that, everything else being equal, scope relations at the Representational Level are reflected in the relative ordering of the corresponding units at the Morphosyntactic Level. That said, the relationship is heavily influenced by the morphosyntactic typology of the language under description. In an isolating language, the relationship is maximally straightforward, with a one-to-one relation between simple words at the Morphosyntactic Level and units at the Representational Level. In an agglutinating language, the same applies, but now to morphemes. Consider the following example from Turkish:

(42) Anlı-y-abıl-ecek-miş-im.
understand-CONN-ABIL-IRR-INFER-1.SG
'I gather I will be able to understand.'
Representational Level: (infer p_i: (ep_i: (irr e_i: [(abil f_i: [(f_j: anlı$_V$(f_j)) ($1x_i$)$_A$] (f_i)) (e_i)$_U$]) (ep_i)) (p_i))
Morphosyntactic Level: (Le_i: (Cl_i: (Vp_i: (Vw_i: [(Vs_i: anlı (Vs_i)) (Aff_i: AbIl (Aff_i)) (Aff_i: EcEk(Aff_i)) (Aff_i: mIş (Aff_i)) (Aff_i: Im (Aff_i))] (Vw_i)) (Vp_i)) (Cl_i)) (Le_i))

In fusional languages, where one affix corresponds to several elements at the Representational Level, the final form cannot be given until the Phonological Level: at the Morphosyntactic Level, we find a placeholder (cf. 15.5.2.7 below). In polysynthetic languages we find little isomorphism between the Representational and Morphosyntactic Levels; the relationship may be further complicated by incorporation (of Words, Phrases or Clauses), as demonstrated in section 15.6.

15.5.2.6 *Relationship between the Representational and Phonological Levels*

Certain features of the Representational Level are realized phonologically. Consider the following example from Scottish Gaelic:

(43) a. Tha an nighean math air bruidhinn.
COP.PRS DEF girl good at talking
'The girl is good at talking.'
b. Tha an nighean mhath air bruidhinn.
COP.PRS DEF girl good ASP talking
'The good girl has been talking.'

In (43a), *math* "good" does not belong to the Individual unit headed by *nighean* "girl" and retains its lexical form /ma/, introduced at the Representational Level; in (43b), however, it functions as a modifier of the feminine head *nighean*, which induces lenition of the first consonant, yielding /ṽa/ at the Phonological Level.

Many ideophones (cf. Voeltz and Kilian-Hatz eds. 2001) exemplify Representational Level units that are transferred directly to the Phonological Level, bypassing the Morphosyntactic Level (since they undergo no morphosyntactic processes). For an example, see section 15.6.

15.5.2.7 *Relationship between the Morphosyntactic and Phonological Levels*

As has been implicit in the preceding discussion, languages differ in whether a particular distinction in Formulation corresponds to effects at the Morphosyntactic or Phonological Level. There appears to be a certain trade-off between the two

Encoding levels, such that a distinction that is encoded at one level need not also be encoded at the other. Thus in Garo (Burling 2004: 67), the "intonation of questions formed with a question word is not much different from the normal statement intonation", but if the final particle *-ma* or *-ni* is omitted, then a rising intonation is required to distinguish the intended Illocution. Rialland and Robert (2001) have shown the non-tone language Wolof not to have any intonational marking of Contrast. The Contrast element is placed in Clause-initial position, followed by a marker (here *laa*) inflected in agreement with the Subject of the following Vp:

(44) Lekkuma mburu mi, ceeb bi laa lekk.
eat.NEG.1SG bread DEF rice DEF CONTR.1SG eat
'I didn't eat the bread, it was the rice I ate.'

This "focus", as the authors call it, "has no effect on the melodic contour of the sentences" (2001: 899).

One important function of the Phonological Level is to provide phonemic form for placeholders introduced at the Morphosyntactic Level. In Spanish, for instance, the placeholder "indpastpf1sg" (corresponding to the interpersonal and representational operators Decl, Past, Perf and a "1sg" argument) appears at the Phonological Level as /e/ in a stressed syllable after verbs of one class and as /i/ after verbs of other classes.

15.6 A WORKED EXAMPLE

Let us now illustrate the four-level analysis outlined above by applying it to a concrete example. The example is from the Australian language Bininj Gun-Wok, more specifically from the Manyallaluk Mayali variety. The example is taken from Bishop and Fletcher (2005: 350), where it is given a prosodic analysis. For the morphosyntactic analysis we rely on Evans's (2003) description of the language:[3]

(45) Ba-rri-ø-yaw-gurrme-ng, wotjbirr.
3.SUBJ(PST)-PL-3.SG.OBJ-child-put-PST.REAL.PF 'smack'
'They put the child down, smack!'

Bininj Gun-Wok has a number of features, illustrated in this sentence, which make it interesting for our purposes, such as its highly synthetic nature, as manifested in the presence of incorporation and crossreference marking, the existence of ideophones used "to represent sounds accompanying actions in the narrative" (Evans 2003: 627), its primative-secundative alignment system, and the occurrence

[3] We are indebted to Nick Evans for his kind help in analyzing this example.

of cumulation in the area of inflection, i.e., the expression of more than one inflectional category in a single morpheme.

Starting our analysis at the Interpersonal Level, we note that the speaker chooses to evoke a single State-of-Affairs in two different Acts, one in which the State-of-Affairs is evoked in terms of a description, and one in which it is evoked in terms of the sound its occurrence provoked. Each of these Acts is Declarative in nature, and the two together constitute a Move. The latter two facts are expressed prosodically, each Declarative Act having a falling contour (Bishop and Fletcher 2005: 335), and the second one falling more pronouncedly, which we assume indicates that it constitutes the closing Act within the Move. The initial Interpersonal Level analysis may thus be given as in (46):

(46) $(M_I{:}\ [(A_I{:}\ [(F_I{:}\ \text{DECL}\ (F_I))\ (P_I)_S\ (P_J)_A\ (C_I)]\ (A_I))$
$(A_J{:}\ [(F_J{:}\ \text{DECL}\ (F_J))\ (P_I)_S\ (P_J)_A\ (C_J)]\ (A_J))$
$](M_I))$

Note that the indices for the Participants do not change, since there is no shift in participant roles, P_I being the Speaker, P_J the Addressee in both acts.

The Communicated Content of A_I consists of an Ascriptive Subact, evoking the property expressed by the verb *gurrme*, and three Referential Subacts, one evoking the entity corresponding to the Actor of the State-of-Affairs designated, and two corresponding to the Undergoer, i.e., two Referential Subacts target a single semantic unit. This latter observation may come as somewhat of a surprise, but reflects the fact that Bininj Gun-Wok is a pronominal argument language. The pronominal prefix on the verb is by itself referential in nature, and need not co-occur with a lexically expressed Undergoer argument, as shown by the optionality of the lexical Undergoer in the following example (Evans 2003: 425–6):

(47) Al-ege daluk gaban-du-ng (bedda).
F-DEM woman 3.SBJ>3.PL.OBJ-scold-NONPST them
'That woman is scolding them.'

This means that in the case of example (45) the Undergoer is referred to twice, in two Referential Subacts, one corresponding to the referential pronominal prefix (in this particular case a zero morpheme), and one to the incorporated argument expression. Note that the incorporated noun *yaw* in (45) must indeed be considered referential, since it would otherwise not be crossreferenced in the portmanteau prefix on the verbal complex. Incorporation can thus be said to be syntactic, not lexical (Smit 2005), and leaves the transitive nature of the verb intact. The Communicated Content of (A_J) contains a single Ascriptive Subact, evoking the sound represented by the ideophone *wotjbirr* that characterizes the State-of-Affairs evoked in (A_I).

Incorporating these observations, we arrive at the more elaborate Interpersonal Level representation in (48):

(48) $(\mathrm{M_I}\!: [(\mathrm{A_I}\!: [(\mathrm{F_I}\!: \mathrm{DECL}\ (\mathrm{F_I}))\ (\mathrm{P_I})_\mathrm{S}\ (\mathrm{P_J})_\mathrm{A}\ (\mathrm{C_I}\!: [(\mathrm{T_I})\ (\mathrm{R_I})\ (\mathrm{R_J})\ (\mathrm{R_K})]\ (\mathrm{C_I}))]$
$(\mathrm{A_I}))\ (\mathrm{A_J}\!: [(\mathrm{F_J}\!: \mathrm{DECL}\ (\mathrm{F_J}))\ (\mathrm{P_I})_\mathrm{S}\ (\mathrm{P_J})_\mathrm{A}\ (\mathrm{C_J}\!: [(\mathrm{T_J})]\ (\mathrm{C_J}))]\ (\mathrm{A_J}))](\mathrm{M_I}))$

Turning now to the Representational Level, the semantic counterpart of $(\mathrm{A_I})$ may be represented as in (49), which additionally shows the alignment with the Interpersonal Subacts discussed above:

(49) Barri-yaw-gurrme-ng.
3.SUBJ.PL(PST)>3.SG.OBJ-child-put-PST.REAL.PF
'They put the child down.'

$(\mathrm{T_I})$ $(\mathrm{R_I})$ $(\mathrm{R_J})$ $(\mathrm{R_K})$

$(\mathrm{p_i}$: (past $\mathrm{ep_i}$: (real $\mathrm{e_i}$: [(pf $\mathrm{f_i}$: [$(\mathrm{f_j}$: /gʊrmϵ/ $(\mathrm{f_j}))$ $(\mathrm{x_i})_\mathrm{A}$ $(\mathrm{x_j}$:
$[(\mathrm{f_k}$: /jaʊ/$(\mathrm{f_k}))\ (\mathrm{x_j})])_\mathrm{U}]\ (\mathrm{f_i}))\ (\mathrm{e_i})])\ (\mathrm{ep_i}))\ (\mathrm{p_i}))$

Within the predication frame at the Representational Level, there is a configuration with a Property $(\mathrm{f_j})$ as the nucleus, and two Individuals $(\mathrm{x_i})$ and $(\mathrm{x_j})$ as the dependents. The Individual $(\mathrm{x_i})$ is not lexically realized but expressed by means of the pronominal prefix. Its identity has to be retrieved from the Contextual Component on the basis of its index. As argued above, the Individual $(\mathrm{x_j})$ is realized twice, once lexically, and once by means of the pronominal prefix. This does not affect the semantic representation, though, just the pragmatic representation.

Semantically speaking, the noun to be incorporated must be a head, since incorporated nouns can take external modifiers in Bininj Gun-Wok, as illustrated in (50) (Evans 2003: 452):

(50) Ga-yau-garrme al-daluk.
3.SBJ>3.SG.OBJ-child-have.NONPST F-woman
'She has a female child.'

The predication frame forms part of a representational frame that shows the hierarchical embedding of the predication frame. The relevant layers shown here are the Propositional Content $(\mathrm{p_i})$, the Episode $(\mathrm{ep_i})$, which carries the absolute tense operator, the State-of-Affairs $(\mathrm{e_i})$, which carries the realis operator, and the Configurational Property $(\mathrm{f_i})$, which carries the perfectivity operator. The fact that these three operators are expressed in a single portmanteau morpheme is a morphosyntactic fact that does not affect their analysis as three different elements at the Representational Level.

In order to formulate the semantic counterpart of $(\mathrm{A_J})$, the status of the ideophone *wotjbirr* should be established. Ideophones have not received systematic treatment in FDG, but what can be said about Bininj Gun-Wok ideophones is that they represent a set of lexical elements that show grammatically distinct behavior and are primarily used for the conventionalized designation of sounds.[4] This justifies setting up a semantic category "S(ound)" for Bininj Gun-Wok. Note that the lexicalized nature of ideophones is reflected, among other things, in the fact that they participate in verbal compounding (Evans 2003: 341).

The semantic counterpart of (A_J) may now be represented as in (51):

(51) wotjbirr
'smack'

$$
\begin{array}{l}
\qquad\quad (T_J) \\
\qquad\qquad | \\
(e_i\text{: [}\quad (s_i\text{: /wɔcbɪr/ }(s_i))\ (e_i)_U])
\end{array}
$$

Note that the index of the State-of-Affairs variable is co-indexed with the one in (49), thus correctly indicating the fact that the ideophone provides an alternative way of characterizing the same event.

The Morphosyntactic Level has the following representation for the counterpart of (A_I):

(52) $(Cl_i\text{: } [(Vw_i)]\ (Cl_i))$

Though the example consists of a single word, we need the clausal layer in (52) so as to allow for the addition of external modifiers. The template for the verbal word is given in (53):

(53) $(Vw_i\text{: } [(Aff_i\text{: /baɪ/}(Aff_i))\ (Ns_i\text{: /jaʊ/ }(Ns_i))\ (Vs_i\text{: /gʊrmɛ/ }(Vs_i))\ (Aff_j\text{: } 138\ (Aff_j))]\ Vw_i))$

The selection of portmanteau pronominal prefixes in Bininj Gun-Wok depends on the way in which the Subject and Object functions are distributed, and is, with third person Subjects only, furthermore dependent on the tense of the verb. The relevance of the Subject function shows up in the fact that there is neutralization of Actor and Undergoer arguments of intransitive predicates, and in the fact that only Subjects can control reflexives and reciprocals, as shown in (54) (Evans 2003: 390):

(54) Barri-marne-ganj-ginje-rr-inj.
3.SUBJ.PL(PST)-BEN-meat-cook-COREF-PST.REAL.PF
'They cooked the meat for each other.'

The relevance of the Object function is apparent in the primative-secundative alignment system of the language, which means that there is neutralization between two-place Undergoers and three-place Recipients and Beneficiaries, as shown in

[4] A more precise characterization suggested to us by Nick Evans would be to say that ideophones in Bininj Gun-Wok denote "synesthetic impressions".

(55) (Evans 2003: 390), in which Object agreement is with the Beneficiary rather than with the Undergoer:

(55) Bandi-marne-ganj-ginje-ng.
3.SUBJ.PL(PST)>3.PL.OBJ-BEN-meat-cook-PST.REAL.PF
'They cooked the meat for them.'

The prefix can thus only be selected after the Subject and Object function have been assigned by the Morphosyntactic Encoder. This is a straightforward process, as there is no true passive voice available in the language (Evans 2003: 574). The tense information necessary for the selection of the appropriate form of the prefix can be retrieved directly from the Representational Level.

The form of the tense suffix is, among other things, dependent on the last syllable of the preceding verbal stem or on the reflexive/reciprocal suffix that may be attached to it (Evans 2003: 323), as is shown contrastively in the examples (54)–(55). This means that the actual form of the suffix can only be selected at the Phonological Level. For this reason a placeholder, here arbitrarily "138", occupies the relevant affix slot.

The ordering of the various components of the verbal word may be represented as in (56):

(56)
P^I P^{F-2} P^{F-1} P^F
(Vw_i: [(Aff_i: /baɪ/ (Aff_i)) (Ns_i: /jaʊ/ (Ns_i)) (Vs_i: /gʊrmɛ/ (Vs_i)) (Aff_j: 138 (Aff_j))] Vw_i))

The morphological possibilities of a Bininj Gun-Wok word are very rich, as shown in Evans (2003: 318), and only partly exploited in the current example. Every verbal word obligatorily contains an initial pronominal complex, usually a portmanteau morpheme, necessarily in the leftmost position, a TMA suffix in the rightmost position, and a (potentially derived) verbal stem immediately preceding the TMA suffix. If there is an incorporated noun, it immediately precedes the (potentially derived) verbal stem. In the process of hierarchical ordering, the TMA suffix is located in P^F. In the process of configurational ordering, the verbal nucleus is placed right in front of the TMA suffix, and the incorporated noun in the next position available left to it. The pronominal portmanteau expression is placed in P^I.

The complex relationships between the Representational and Morphosyntactic Levels may now be shown as in (57):

(57) (past ep_i: (real e_i: [(pf f_i: [(f_j: /gʊrmɛ/ (f_j)) $(x_i)_A$ (x_j: [(f_k: /jaʊ/ (f_k)) $(x_j)_Ø$])$_U$] (f_i)) $(e_i)_Ø$]) (ep_i))

(Vw_i: [(Aff_i: /baɪ/ (Aff_i)) (Ns_i: /jaʊ/ (Ns_i)) (Vs_i: /gʊrmɛ/ (Vs_i)) (Aff_j: 138 (Aff_j))] Vw_i))

What this representation shows is that there is a one-to-one relation between lexical elements at the Representational level and stem slots within the morphosyntactic word template, but a many-to-one relationship between non-lexical elements at the Representational level and affix slots at the Morphosyntactic level, the latter point clearly showing the cumulative nature of the inflectional affixes in Bininj Gun-Wok. It furthermore shows that independent units at the semantic level enter into the internal constitution of a single Morphosyntactic word.

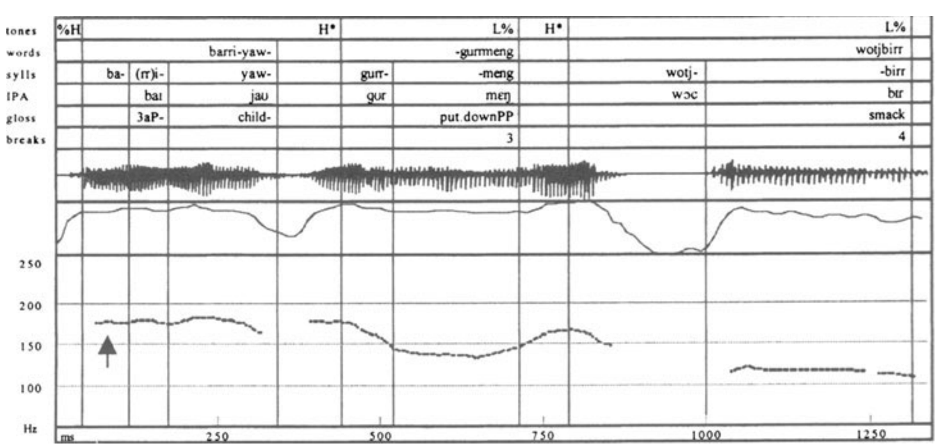
The morphosyntactic counterpart of A_J is straightforward:

(58) $(\mathrm{Iw_i}: (\mathrm{Is_i}:$ /wɔcbɪr/ $(\mathrm{Is_i}))\ (\mathrm{Iw_i}))$

Given that the class of ideophones constitutes a special word class in Bininj Gun-Wok, we use the category (Iw) to account for them. This is furthermore a good example of an Act corresponding to a single word, i.e., a holophrastic expression.

The formalization at the Phonological Level of example (45) is as in (59). The spectrogram, shown below the formula, is taken from Bishop and Fletcher (2005: 350), on which our argumentation is based.

(59) (f $\mathrm{U_i}$: [(f $\mathrm{IP_i}$: [(f $\mathrm{PP_i}$: [(h $\mathrm{F_i}$: [(s $\mathrm{s_i}$: /baɪ/ $(\mathrm{s_i})$) $(\mathrm{s_j}$: /jaʊ/ $(\mathrm{s_j})$)] $(\mathrm{F_i})$) $(\mathrm{F_j}$: [(s $\mathrm{s_k}$: /gʊr/ $(\mathrm{s_k})$) $(\mathrm{s_l}$: /mɛŋ/ $(\mathrm{s_l})$)] $(\mathrm{F_j})$)] $(\mathrm{PP_i})$) $(\mathrm{IP_i})$) (f $\mathrm{IP_j}$: (f $\mathrm{PP_j}$: [$(\mathrm{F_k}$: [(s $\mathrm{s_m}$: /wɔc/ $(\mathrm{s_m})$) $(\mathrm{s_n}$: /bɪr/ $(\mathrm{s_n})$)] $(\mathrm{F_i})$) $(\mathrm{PP_j})$)] $(\mathrm{IP_j})$)] $(\mathrm{U_i})$



Bishop and Fletcher (2005: 358) indicate that the pause between *barriyawgurrmeng* and *wotjbirr* has the index 3, indicating an utterance-medial break between intonational phrases, hence the two IPs within the Utterance $\mathrm{U_i}$ (for further detail on break indices, see Bishop and Fletcher 2005: 352–4). Each of the IPs corresponds to an Act at the Interpersonal Level. That there is one Utterance (U) here is supported by the final pause with break index 4, which indicates an utterance boundary (Bishop and Fletcher 2005: 358). The Utterance as a whole corresponds to a Move at the Interpersonal Level.

Both IPs have a falling contour, as shown by their f-operators, which expresses their Declarative nature. The recording shows that the second IP has a particularly clear fall from /wɔc/ to /bɪr/, which we interpret as a paratone effect indicative of the end of a Move, and indicated by an f-operator on U_i. The first IP has a high initial foot (h F_i), reflecting Bishop and Fletcher's (2005: 350) identification of this example as having, in their terms, an "initial high boundary tone (%H)".

In the example each IP contains one PP, and in another Bininj Gun-Wok dialect it is known that PPs have a falling contour ("tonally marked (with a low tone) at its right edge", Bishop and Fletcher 2005: 341). Bishop and Fletcher surmise that this is true of the Manyallaluk Mayali dialect, too. This is indicated by the f-operator on the PPs.

The Bininj Gun-Wok IP has a single nuclear accent and the "boundary tone" is signaled on the last or the penultimate and last syllables of the IP (Bishop and Fletcher 2005: 342). This is reflected in the analysis: each IP has a falling operator, which the Articulator will attribute to the final stressed syllable of each, namely /gʊr/ and /wɔc/. The following syllables /mɛŋ/ and /bɪr/ are correspondingly produced at a lower pitch than the preceding syllables.

The level of the Phonological Word (PW) has not been found necessary for a description of Bininj Gun-Wok intonation (Bishop and Fletcher 2005: 339) and has accordingly not been included here.

In each Foot (F), it is the first syllable that is stressed (as always in Bininj Gun-Wok; Bishop and Fletcher 2005: 340 point out that the foot is "trochaic and unbounded", giving an example of a foot with three unaccented syllables: *gorlomomo* "fresh water crocodile"). This is indicated by means of the s-operators on the first syllables within each foot. Foot F_i, which corresponds to the morphosyntactic structure *barriyaw* could alternatively be analyzed as having three syllables, again with stress on the first. In this case s_i in (59) would be expanded as in (60):

(60) (s s_i: /ba/ (s_i)) (s_{i+1}: /rɪ/ (s_{i+1}))

Then the collapsing of the first two into one syllable in the actual realization of the utterance would have to be left to the Articulator, possibly as a feature of *allegro* speech.

Of the various prosodic features manifested by this example, we can thus say that those that apply at the level of the Utterance and the Intonational Phrases are meaningful, in the sense that they express distinctions relevant at the Interpersonal Level, while the other ones correspond to default settings. The overall correspondences between the Interpersonal Level and the Phonological Level may then be indicated as in (61), where we leave out the intermediate Representational and Morphosyntactic Levels for the sake of readability:

(61) $(\mathrm{M_I}: \quad [\quad (\mathrm{A_I}:[(\mathrm{F_I}: \mathrm{DECL}\ (\mathrm{F_I}))....]\ (\mathrm{A_I})) \quad (\mathrm{A_J}: [(\mathrm{F_J}: \mathrm{DECL}\ (\mathrm{F_J}))....]\ (\mathrm{A_J}))\ ]\ (\mathrm{M_I}))$

$(f\ \mathrm{U_i}: \quad [\quad (f\,\mathrm{IP_i}) \quad (f\,\mathrm{IP_j}) \quad](\mathrm{U_i})$

15.7 Dynamic variation

Our argument in the preceding sections was based on static synchronic data, from both a language-specific and a typological perspective. FDG, however, also aims to offer a framework for the analysis of dynamic data, be these concerned with language acquisition and language loss, or language genesis and language change. We can only touch upon these issues briefly here.

The FDG framework offers, as noted by Butler and Taverniers (2008), two major predictions as regards dynamic variation: one concerns the variational step between the Representational and the Interpersonal Levels, the other the variational steps between the various hierarchically ordered Layers at each Level. In both cases, the actual manifestation of variational steps will be at the Morphosyntactic and Phonological Levels.

As an example of the step between the Interpersonal Level and the Representational Level, consider the status of adverbial conjunctions in English. Hengeveld and Wanders (2007) show that a basic distinction can be made in English between lexical and grammatical conjunctions: the former can be modified, while the latter cannot, as shown in (62)–(63):

(62) He arrived **three hours** before she left.

(63) *He continued walking around **three hours** until the meeting began.

Both types of conjunction do, however, admit modifiers that have scope over the entire conjunctional phrase, as shown in (64)–(65):

(64) He arrived **exactly** three hours before she left.

(65) He continued walking around **exactly** until the meeting began.

The conjunctions shown in (62)–(65) are all operative at the Representational Level. Conjunctions at the Interpersonal Level admit neither type of modification, as illustrated in (66):

(66) *He is a nice guy, **exactly** while she is a rather unpleasant character.

This point is particularly evident in cases in which a conjunction can be used at both Levels, as in (67)–(68):

(67) *Watch out, **exactly because** there is a bull in the field.

(68) Providing food assistance is not easy **exactly because** the infrastructure is lacking.

In (67) the causal clause motivates a discourse act at the Interpersonal Level, but in (68) it provides the reason for a State-of-Affairs at the Representational Level. From facts like these one may derive the conclusion that the grammaticalization of conjunctions goes hand in hand with their developing an interpersonal function from their originally representational function.

The variational steps between the various hierarchically ordered Layers at one specific level may be illustrated by the acquisition of operator categories at the Representational Level. Drawing on a wealth of data from English child language acquisition, Boland (2006) shows that operators from lower layers are acquired earlier and/or more rapidly than operators from higher layers. More specifically, she shows that aspectual operators (Property layer f) are acquired more rapidly and earlier than tense operators (State-of-Affairs layer e), which in turn are acquired before operators expressing a propositional attitude (Propositional Content layer p). She furthermore shows that this observation holds for the acquisition of a wide range of typologically different languages, and is paralleled by the diachronic developments in this domain.

15.8 Conclusion

This concludes our overview of the FDG model, a structural-functional theory of language with a strong typological basis. Its defining characteristics can be summarized as follows: (i) FDG has a top-down organization; (ii) FDG takes the Discourse Act rather than the sentence as its basic unit of analysis; (iii) FDG is designed as the grammatical component of a wider theory of verbal interaction in which it is connected to Conceptual, Contextual, and Output Components; (iv) FDG contains Interpersonal, Representational, Morphosyntactic, and Phonological Levels. This architecture is applied to both static and dynamic data.

CHAPTER 16

GRAMMATICALIZATION AND LINGUISTIC ANALYSIS

BERND HEINE
HEIKO NARROG

16.1 What is grammaticalization?

Grammaticalization can be understood as a process by which grammar is created (Croft 2006: 366); more narrowly, we define it as the development from lexical to grammatical forms, and from grammatical to even more grammatical forms.[1] Since the development of grammatical forms is not independent of the constructions to which they belong, the study of grammaticalization is in the same way concerned with constructions, and with even larger discourse segments (see Traugott and Heine 1991*a*; 1991*b*; Heine, Claudi and Hünnemeyer 1991; Hopper and Traugott 2003; Bybee, Perkins and Pagliuca 1994; Lehmann 1982 [1995]; Heine 1997*b*; Kuteva 2001; Heine and Kuteva 2002; 2007, section 1.2). Grammaticalization theory is a theory to the extent that it offers an explanatory account of how and why grammatical categories arise and develop, and why they are structured the way they are.

Consider the example in (1). The two sentences have a number of properties in common, such as the verbal item *use* constructed in the past tense. But there are also

[1] For a fairly comprehensive list of definitions that have been proposed for grammaticalization, see Campbell and Janda (2001).

differences; for example, the item *used* has a nominal complement (*all the money*) and a lexical meaning in (1a), while in (1b) it has an infinitival verb (*to come*) and expresses the grammatical function of a durative aspect in (1b). Grammaticalization theory aims at providing explanations for similarities and differences of such constructions; we will return to this example in section 16.5.

(1) English
  a. Wilson used all the money.
  b. Gretchen used to come on Tuesdays.

A number of alternative approaches have been proposed to deal with grammaticalization phenomena. Some of these approaches highlight the semantic aspects of the process (e.g., Heine et al. 1991; Heine 2003), others focus on pragmatic aspects (e.g., Hopper and Traugott 2003; Traugott and Dasher 2002) or discourse considerations (Givón 1979; Hopper 2001), and still others on grammaticalization as the result of repetition on language processing, thus highlighting morphological and phonological aspects (e.g., Bybee 2003; Bybee 2006*a*; Bybee, this volume). More recently, the concept of grammaticalization has been extended to the analysis of syntax (e.g., Roberts and Roussou 2003; van Gelderen 2004). In the framework proposed here, all these different aspects play some role. This framework rests on the assumption that the main motivation underlying grammaticalization is to communicate successfully. To this end, one salient human strategy consists in using linguistic forms for meanings that are concrete, easily accessible, and/or clearly delineated to also express less concrete, less easily accessible and less clearly delineated meanings. To this end, lexical or less grammaticalized linguistic expressions are pressed into service for the expression of more grammaticalized functions.

Grammaticalization theory is concerned with regularities in language use as they can be observed in spoken and written linguistic discourse on the one hand and in language change on the other. It does not presuppose any formal theoretical constructs, such as a distinction between an E-language and an I-language, nor does it require assumptions to the effect that "language"—however this notion may be defined—is or should be conceived of as a system.

16.2 Theoretical assumptions

For grammaticalization theory, it is an important assumption that change leading to the genesis of grammatical material in natural languages is not random but takes place along certain paths. These paths are crosslinguistically replicable and exhibit a specific directionality, most importantly the directionality from lexical to grammatical, and from grammatical to more grammatical structures, which is defining for grammaticalization theory. Thus, grammaticalization is essentially conceived as a unidirectional process. On the other hand, examples contradicting the unidirectionality principle have been proposed (see especially Campbell 1991; Ramat 1992; Frajzyngier 1996; and especially Newmeyer 1998: 260ff.), usually referred to as instances of "degrammaticalization", but more appropriately to be called "antigrammaticalization" (Haspelmath 2004*b*). A number of these examples, however, are controversial or subsequent research shows that they do not contradict the unidirectionality principle (see, for example, Haspelmath 2004*b*; Andersson 2006). Note further that so far no instance of a complete reversal of a grammaticalization process has been documented (Newmeyer 1998: 263). While we lack reliable statistics, two authors have come up with specific estimates on the relative frequency of grammaticalization vis-à-vis other processes. Newmeyer (1998: 275–6, 278) claims that only about 90% of grammatical changes are in accordance with the unidirectionality principle (that is, they are "downgradings" in his terminology), and Haspelmath (1999*a*: 1,046) suggests that about 99% of all shifts along the lexical/functional continuum are grammaticalizations. On a conservative estimate then, at most one tenth of all grammatical developments can be suspected to be counterexamples to the unidirectionality principle.

Three different stances have been taken to deal with "degrammaticalization". First, it has been argued that since there are some cases of "degrammaticalization", the unidirectionality hypothesis is false (Campbell 2001). Second, it has been suggested that this hypothesis is largely though not entirely true; it takes care of a robust tendency of grammatical change (Haspelmath 1999*a*; Hopper and Traugott 2003). And, third, a number of scholars hold the opinion that the hypothesis is true and cases of presumed "degrammaticalization" can be accounted for by means of alternative principles. Principles that have been invoked are, on the one hand, morphosyntactic, like exemplar-based analogical change (Kiparsky 2005), and, on the other hand, cognitive and communicative forces, such as euphemism, exaptation, and adaptation (Heine 2003; Narrog 2004; 2007).

The relationship between grammaticalization and lexicalization has drawn the particular attention of researchers more recently. Brinton and Traugott (2005) argue that grammaticalization and lexicalization are complementary processes, which exhibit similarities but also clear differences. For example, decategorialization, bleaching, subjectification, and increased productivity and frequency are typical for grammaticalizations but not for lexicalizations (cf. Brinton and Traugott 2005: 104–10).

A different issue, pertaining to the theoretical status of grammaticalization, is the relationship between grammaticalization, on the one hand, and processes of language change that have been traditionally identified in historical linguistics, on the other hand, in particular reanalysis and analogy (cf. Joseph 2001).

Grammaticalization and reanalysis are clearly not identifiable with each other. Grammaticalization arguably always involves reanalysis (for a counter view, see Haspelmath 1998), but reanalysis is also involved in changes that have little or nothing to do with grammaticalization, as for example compounding, derivation, or word order change, and even in changes that apparently run counter to grammaticalization (cf. Narrog 2007). Something similar can be said for analogy. Analogy is generally involved in the spread of patterns in a linguistic system and is not confined to grammaticalization. However, grammaticalization arguably needs a stage of analogical spread of the innovated pattern(s) across the linguistic system (cf. Hopper and Traugott 2003, ch. 3).

16.3 Methodology

Grammatical change has been described in terms of a wide variety of different models. In functional research on grammaticalization, which probably amounts to more than 90% of the work in this area, emphasis has been on two aspects of change. One concerns semantics, in that this process is first and foremost one that leads from less grammaticalized to more grammatical meanings. The second aspect concerns pragmatics, and in particular the role of context. In the standard works treatments mentioned above, attempts are made to reconcile these two aspects in some way or other. The methodology employed here (see also Heine and Kuteva 2002; 2007, section 1.2) rests on the assumption that grammaticalization is based on the interaction of pragmatic, semantic, morphosyntactic, and phonetic factors. There is a wide range of criteria that have been proposed to describe grammaticalization (see, for example, Lehmann 1982 [1995]; Heine et al. 1991; Hopper and Traugott 2003; Bybee et al. 1994); in the model proposed here are the four parameters listed in (2). Alternative theoretical concepts that have been proposed are, among others, syntacticization, morphologization, obligatorification,[2] subjectification, etc. It is argued that they can be accounted for essentially with reference to these four parameters. Henceforth these parameters will be used as a tool for identifying and describing instances of grammaticalization.

[2] Some students of this paradigm of linguistics argue that obligatorification, whereby the use of linguistic structures becomes increasingly more obligatory in the process of grammaticalization, should be taken as a definitional property of this process. As important as obligatorification is (see Lehrmann 1982 [1995]), it is neither a *sine qua non* for grammaticalization to take place, nor is it restricted to this process, occurring also in other kinds of linguistic change, such as lexicalization. Within the present framework, obligatorification—as far as it relates to grammaticalization—is a predictable by-product of decategorialization.

(2) Parameters of grammaticalization[3]
  a. Extension, i.e., the rise of new grammatical meanings when linguistic expressions are extended to new contexts (context-induced reinterpretation).
  b. Desemanticization (or "semantic bleaching"), i.e., loss (or generalization) in meaning content.
  c. Decategorialization, i.e., loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms.
  d. Erosion ("phonetic reduction"), i.e., loss in phonetic substance.

Each of these parameters concerns a different aspect of language structure or language use; (2a) is pragmatic in nature, (2b) relates to semantics, (2c) to morphosyntax, and (2d) to phonetics. Except for (2a), these parameters involve loss in properties. But the process cannot be reduced to one of structural "degeneration" in that there are also gains. Some authors have stressed that, while linguistic items undergoing grammaticalization lose in semantic, morphosyntactic, and phonetic substance, they also gain in properties characteristic of their uses in new contexts (e.g., Hopper and Traugott 2003). As a result of the acquisiton of new properties, in some cases their meaning and syntactic functions may show little resemblance to their original use.

The ordering of these parameters reflects the diachronic sequence in which they typically apply: Grammaticalization tends to start out with extension, which triggers desemanticization, and subsequently decategorialization and erosion. Erosion is the last parameter to be involved; as we will see below, in many of the examples to be presented, erosion is not (or not yet) a relevant parameter.

16.3.1 Extension

Of all the parameters, extension is the most complex one. It has a sociolinguistic, a text-pragmatic, and a semantic component. The sociolinguistic component concerns the fact that grammaticalization starts with innovation (or activation) as an individual act, whereby some speaker (or a small group of speakers) proposes a new use for an existing form or construction, which is subsequently adopted by other speakers, ideally diffusing throughout an entire speech community (propagation; see, for example, Croft 2000: 4–5). The text-pragmatic component involves the extension from a usual context to a new context or set of contexts, and the gradual spread to more general paradigms of contexts. The semantic component finally

[3] The use of the term "parameter" must not be confused with that found in some formal models of linguistics. For an alternative account of grammaticalization within a Chomskyan Minimalist framework, see van Gelderen (2004), where this process is described in terms of economy principles, entailing in particular a syntactic shift from specifier to head, e.g., from main verb to auxiliary, from demonstrative to definite article, etc.

leads from an existing meaning to another meaning that is evoked or supported by the new context. Thus, text-pragmatic and semantic extension are complementary aspects of one and the same general process characterizing the emergence of new grammatical structures.

It has been argued or implied that the main trigger of grammaticalization is frequency of use: The more often a given form or construction occurs, the more likely it is that it will reduce in structure and meaning and assume a grammatical function (Bybee 1985; 2006*a*; Krug 2003; Diessel 2005: 24; see especially the contributions in Bybee and Hopper 2001). In fact, extension to new (sets of) contexts implies a higher rate of occurrence of the items concerned, and words more probable in a specific context are more likely to be reduced than less probable ones (cf. the Probabilistic Reduction Hypothesis of Jurafsky et al. 2001). Furthermore, when a grammatical item whose use is optional is used more frequently, its use may become obligatory. Nevertheless, we have found neither compelling evidence to support the hypothesis that frequency is immediately responsible for grammaticalization, nor that grammatical forms are generally used more frequently than their corresponding less grammaticalized cognates (see Heine and Kuteva 2007, section 1.2).

16.3.2 Desemanticization

Desemanticization is an immediate consequence of extension: Use of a linguistic expression E in a new context C entails that E loses part of its meaning that is incompatible with C. It is frequently triggered by metaphoric processes (Lakoff and Johnson 1980; Lakoff 1987). For example, a paradigm case of grammaticalization involves a process whereby body part terms ("back", "head", etc.) are reinterpreted as locative adpositions ("behind", "on top of") in specific contexts (cf. section 16.5). Via metaphoric transfer, concepts from the domain of physical objects (body parts) are used as vehicles to express concepts of the domain of spatial orientation, while desemanticization has the effect that the concrete meaning of the body parts is bleached out, giving way to some spatial schema. In a similar fashion, when an action verb (e.g., English *keep, use, go to*) is reinterpreted as a tense or aspect auxiliary (see section 16.4), this can be understood to involve a metaphoric process whereby a concept of the domain of physical actions is transferred to the more abstract domain of temporal and aspectual relations.[4] Once again, this leads to the desemanticization of lexical meaning, namely that of the action verbs (Heine et al. 1991; Heine 1997).

[4] This is a simplified rendering of the process concerned; for a detailed account of such a process, see Heine and Miyashita (2008).

16.3.3 Decategorialization

Once a linguistic expression has been desemanticized it tends to lose morphological and syntactic properties characterizing its earlier use but being no longer relevant to its new use. Decategorialization entails in particular the changes listed in (3):

(3) Salient properties of decategorialization
   a. Loss of ability to be inflected.
   b. Loss of the ability to take on derivational morphology.
   c. Loss of ability to take modifiers.
   d. Loss of independence as an autonomous form, increasing dependence on some other constituent.
   e. Loss of syntactic freedom, e.g., of the ability to be moved around in the sentence in ways that are characteristic of the non-grammaticalized source item.
   f. Loss of ability to be referred to anaphorically.
   g. Loss of members belonging to the same grammatical paradigm.

In accordance with this list, nouns undergoing decategorialization tend to lose morphological distinctions of number, gender, case, etc., the ability to combine with adjectives, determiners, etc., to be headed by adpositions; they lose the syntactic freedom of lexical nouns, and the ability to act as referential units of discourse. In a similar fashion, when a demonstrative develops into a clause subordinator, as has happened in many languages of the world (see Heine and Kuteva 2007, ch. 5), it loses salient categorical properties. Verbs undergoing decategorialization tend to lose their ability to be inflected for tense, aspect, negation, etc., to be morphologically derived, to be modified by adverbs, to take auxiliaries, to be moved around in the sentence like lexical verbs, to conjoin with other verbs, to function as predicates, and to be referred to, for example, by pro-verbs. Finally, they lose most members of the grammatical paradigm to which they belong by changing from open-class items to closed-class items.

16.3.4 Erosion

Erosion means that in the course of or as a result of undergoing grammaticalization, a linguistic expression loses in phonetic substance. Note that erosion is not confined to grammaticalization but is a more general process in language change that may also affect lexical items (cf. Hock 1991*b*, ch. 5). However, erosion is linked to high-frequency use (cf. Bybee 2003), and thus it is a corollary of grammaticalization which implies increase in usage frequency (Bybee and Beckner, this volume). Therefore, it usually occurs at a later stage in the grammaticalization process and is by no means a requirement for grammaticalization to happen. Erosion can be of two kinds. First, it may involve entire morphological units. Thus, when the Old English phrase *Þa hwile Þe* "that time that", or any of its variants, was grammaticalized to the temporal and concessive subordinator *while* in Modern English, this meant that morphological segments were lost, and much the same happened in the case of its Old High German counterpart *al di wila daz* "all the time that", which was grammaticalized to the causal subordinator *weil* "because" in Present-Day German, which is also characterized by loss of morphological elements. Such cases are referred to as morphological erosion (Heine and Reh 1984). More commonly, however, change is restricted to phonetic erosion, that is, to phonetic properties, in particular the ones listed in (4), or any combination thereof.

(4) Kinds of phonetic erosion
 a. Loss of phonetic segments, including loss of full syllables.
 b. Loss of suprasegmental properties, such as stress, tone, or intonation.
 c. Loss of phonetic autonomy and adaptation to adjacent phonetic units.
 d. Phonetic simplification.

In quite a number of cases, both morphological and phonetic erosion tend to be involved. For example, the grammaticalization of the phrase *by the side of* to the preposition *beside* in Modern English, or of *by cause of* to *because (of)* appear to have involved both morphological and phonetic erosion. Similarly, the development of the Latin phrase *(in) casa* "in the house (of)" via Old French *(en) chies* "at" to the Modern French locative preposition *chez* "at" involved loss of both morphological and phonological substance.

16.4 Overlapping categories

In most models of contemporary linguistic analysis, grammatical phenomena are described in terms "classical categories", that is, taxonomic units having the properties listed in (5).

(5) Properties of classical categories (Taylor 1989: 23–4)
 a. They are defined in terms of necessary and sufficient features.
 b. Features are binary.
 c. The categories have clear boundaries.
 d. All members of the category have equal status.

In spite of all the advantages that such categories have over alternative taxonomic constructs, they are insufficient for a satisfactory description and understanding of the nature of numerous linguistic phenomena, and therefore some linguists have turned to alternative taxonomic models, in particular to prototype models, Wittgensteinian family-resemblance models, or an exemplar model (Bybee and Beckner, this volume). Students of grammaticalization have drawn attention to the gradient nature of grammatical phenomena and have proposed continuum models of categorization, such as grammaticalization chains (Heine 1992) or clines (Hopper and Traugott 2003). What is highlighted in the latter models is that grammaticalization generally proceeds along the same kind of stages and hence can be accounted for best in terms of a three-stage model of the kind represented in (6).

(6) The overlap model (Heine 1993: 48–53)
   I There is a linguistic structure A.
   II A acquires a second structure B in specific contexts (= A/B).
   III In some other context, A is lost, with the effect that there is only B.

The result of this process is that there is a chain-like structure A, A/B, B (henceforth referred to as I/II/III structure), in that the different stages of the process surface as contextually defined variants in the synchronic form of a language. Due to specific historical processes that may happen, however, there are three main variants of this model. One concerns situations where the process is arrested at stage II, in which case there is no structure B, the resulting structures being A and A/B (= I/II structures). Alternatively, A (= stage I) can be lost, in which case the remaining structures are A/B and B (= II/III structures), or else A/B is lost, resulting in the presence of the structures A and B (= I/III structures).

Furthermore, there are the following additional possibilities, both crosslinguistically widely attested. Either the process never takes place, that is, there is no change from I to II, with the result that there is only structure A. Alternatively, both A and A/B are lost, with the effect that there is only structure B. We will look no further at these two situations because they do not offer any problems for synchronic grammatical taxonomy.

The situation is different in the case of I/II/III structures; all languages that we are familiar with have such structures. We may illustrate the nature of I/II/III structures with a crosslinguistically widely attested example concerning the marking of reflexive and reciprocal categories. Most languages have conventionalized constructions for expressing reflexive and reciprocal concepts. That these are two distinct concepts is suggested, for example, by the fact that many languages use different grammatical forms for the two, cf. English *They hate themselves* (reflexive) vs. *They hate each other* (reciprocal). In yet other languages the two are not formally distinguished, forming what Heine and Miyashita (2008) call a REF-REC category. This category has most or all of the properties listed in (7).

(7) Properties of REF-REC categories
   a. With singular antecedent referents, the category expresses reflexivity only (= A).

b. With multiple antecedents (i.e., plural or conjoined subject referents), the category is likely to be ambiguous, expressing both reflexivity and reciprocity (= A/B).[5]
c. With multiple antecedents of certain verbs (i.e., "symmetrical" or "inherently reciprocal" verbs) or in specific other contexts, the category expresses reciprocity only (B).
d. In view of their overlapping structure (cf. (b)), there is no categorial boundary setting reflexive and reciprocal readings apart.
e. Accordingly, essentially the same syntactic construction is employed for both reflexive and reciprocal functions.
f. Compared to reciprocals which are not part of REF-REC categories, the reciprocal use does not exhibit any high degree functional variation.
g. In the relevant literature, the reflexive meaning tends to be portrayed as the basic one, or as being more basic than the reciprocal one.
h. The reflexive meaning is the unmarked one in the sense that it is less constrained in its use; for example, it is associated with both singular and multiple antecedents whereas the reciprocal is restricted to multiple antecedents.
i. REF-REC categories are the result of a historical process leading from reflexive to reciprocal meaning.

The following example from German illustrates the structure of such an overlapping category:[6] (a) shows the reflexive-only structure (= A) used with singular subject referents, (b) the ambiguity structure (= A/B), and (c) the reciprocal-only structure (= B).

(8) German

a. Er wusch sich.
he wash.PAST REF
'He washed (himself).' (A)

b. Sie wuschen sich.
they wash.PAST.PL REF
i. 'They washed themselves.' (A)
ii. 'They washed each other.' (B)

c. Sie küssten sich.
they kissed.PAST.PL REF
'They kissed (each other).' (B)

The German category conforms in every respect to the characterization in (7). That the reflexive use is less constrained than the reciprocal one can be seen in the fact

[5] There are a few languages, such as Lithuanian (Maslova and Nedjalkov 2005: 430), where there is no ambiguity in that only one of the two meanings is possible for a given verb.

[6] The German category takes a suppletive form. With third person referents the shape is *sich* while with first and second person referents appropriate personal pronouns are used.

that the latter is more strongly grammaticalized, having lost the ability to be stressed (erosion). Thus, *sich* in (8a) can be stressed or unstressed, but when *sich* is stressed in (8b), the reciprocal meaning is suppressed.

REF-REC categories are not strictly a rare phenomenon. In their corpus of roughly 150 languages from all major regions of the world, Heine and Miyashita (2008) found that at least every third language has such a category.[7] The data available suggest that the rise of REF-REC categories is the result of a gradual transition from reflexive to reciprocal uses, while a process in the opposite direction is unlikely to happen. The initial stage of this process is provided by reflexive-only categories (stage I). In the intermediate stage II, the reflexive category does not rule out a reciprocal interpretation under certain circumstances. Such a situation can be generalized, in that reciprocity becomes a regularly distinguished reading of the reflexive category, and once its use is extended to inherently reciprocal ("symmetric") verbs, there is a reciprocal reading only (stage III). Accordingly, the REF-REC category of German and hundreds of other languages takes the form of a I/II/III structure.

To conclude, the presence of this type of overlapping category in a wide range of languages is the result of a process whereby three different stages of grammaticalization surface in the synchronic state of the languages concerned in the form of three different contextually defined use patterns.

16.5 Lexical-functional splits—the case of auxiliaries

We observed in the preceding section that grammaticalization processes can have a number of consequences for synchronic categorization, one possible consequence being the rise of overlapping I/II/III categories. A well-know example for this process is the English *be going-to* construction (Pérez 1990; Bybee et al. 1994; Langacker 1998*a*). In Middle English there was only a lexical stage, at which the verb *go* expressed spatial motion. Toward the end of the 15th century, the auxiliary use with an intention and subsequently the temporal function started to emerge. Subsequently, there are examples which are suggestive of the temporal function.

[7] Roughly the same finding is made by Maslova and Nedjalkov (2005: 430–3). As the data provided by these authors show, languages distinguishing between reflexive and reciprocal constructions constitute clearly the majority of the world's languages (72.3%) while languages with REF-REC categories form a significant minority (27.7%, i.e., 44 out of 159 languages); however, if one adds the 16 languages of their sample which have a REF-REC construction in addition to a distinct reciprocal construction, then there are altogether 34.3% (60 out of 175 languages) having a REF-REC category.

In present-day English, the temporal meaning has largely been generalized with non-finite (i.e., infinitival) verbs. However, examples that are ambiguous between lexical and auxiliary uses are still available when the complement describes a volitional action and does not contradict the motion encoded in *go*. Accordingly, in present-day English we find both the lexical category (stage I) in (9a), the ambiguous transitory stage (II) in (9b), and the unambiguous auxiliary (stage III) in (9c) (cf. Langacker 1998*a*: 79):[8]

(9) English
  a. Mary **is going to** town.
  b. Mary **is going to** buy vegetables.
  c. Mary **is going to** come soon.

Among the alternative consequences that we mentioned above there is one where the intermediate stage II is lost, with the result that there is a I/III structure. This is crosslinguistically a highly common process, leading to a split where I and III are contrasting structures that nevertheless share a number of similarities as a result of their common origin. The grammaticalization from lexical verbs to what we refer to loosely as auxiliaries provides a paradigm example of such a process.

A crosslinguistically common situation is one where one and the same linguistic form has two contrasting meanings each associated with a different construction, where one meaning is that of a lexical verb while the other is that of an auxiliary typically expressing functions of tense, aspect, or modality. English examples of such doublets can be found in (10), where (10a) is suggestive of the lexical stage I and (10b) of functional meanings and constructions of stage III, and where the items printed in bold in (10a) are main verbs and the corresponding ones in (10b) auxiliaries for grammatical concepts, i.e., concepts for durative (*kept*) or habitual aspect (*used to*), or for deontic modality (*has to*).

(10) English
  a. He **kept** all the money. He **used** all the money. He **has** all the money.
  b. He **kept** complaining. He **used** to come on Tuesdays. He **has** to pay.

In the case of *have to*, for example, ambiguous constructions of the type *I have a letter to mail* (cf. Heine 1993: 41f.; Krug 2000: 54–60) are no longer available in present-day English, leading to a loss of the transitory stage II.

The relevant literature is rife with discussions of such examples, which we summarily classify as manifestations of an auxiliation process. In accordance with the overlap model proposed in (6), this process can be sketched as in (11) (see Heine 1993; Bybee et al. 1994; Kuteva 2001 for discussion).

[8] If *going to* is fused to *gonna*, the ambiguity with the lexical use is resolved, since the lexical item *going* does not fuse with *to*.

(11) Salient characteristics of auxiliation

I There is a lexical verb (V_1) taking a non-finite verb as a complement (C).

II In specific contexts, V_1 can be interpreted alternatively as a lexical verb or as expressing a grammatical function relating to tense, aspect, or modality, and C has either the function of a complement or of a new main verb.

III In certain other contexts, V_1 expresses exclusively the grammatical function, and C that of the new main verb.

IV The use of the grammatical function is generalized, and C can only be understood as the main verb whenever there is a non-finite verb as a complement (= loss of stage II).

Auxiliation thus leads to a split between a lexical and a functional category, in that in addition to construction (12a) there is also now (12b).

(12) A morphosyntactic skeleton of lexical and corresponding auxiliary constructions[9]

a. main verb (V_1)–non-finite verb complement (C)	Stage I
b. auxiliary (A)–main verb (V_2)	Stage III

Still, (12a) and (12b) share a number of properties. First, the morphological form (*have, use* etc.) can be the same (or at least similar). Second, like V_1, A can exhibit properties of a lexical main verb, such as being inflected for person and number agreement, taking tense and negation markers, etc., while V_2 is non-finite, taking the form of an infinitival, participial, or gerundival word. And, third, in spite of the change from (12a) to (12b), the latter tends to retain relics of its erstwhile verb–complement syntax.

But in spite of such commonalities there are a number of systematic differences between (12a) and (12b), in particular the ones summarized in Table 16.1. These differences are the result of the grammaticalization from main verb to auxiliary construction and can be described by means of the parameters of grammaticalization that we presented in (2):

(13) From lexical to auxiliary construction

a. V_1 loses most or all of its lexical semantics, including the ability to control an action (desemanticization). And in much the same way as it loses in semantic properties, V_1 acquires properties of a functional marker.

b. V_1 loses salient morphosyntactic properties of a verb, such as the ability to select the subject or to take adjuncts, or to be associated with the whole range of tense, aspect, and modality marking (decategorialization).

c. V_1 tends to be phonetically reduced, it may lose the ability to be stressed (erosion).

[9] In languages having a verb-final syntax, the order of the two constituents is reversed. Note that while (12) and (13) capture common characteristics of auxiliaries in a number of languages, there are also other processes of auxiliation that we cannot take into account in the present chapter.

Table 16.1. Contrasting properties of lexical and their corresponding auxiliary constructions

	Lexical category	Functional category
Syntactic structure	V_1–C	A–V_2
Meaning	V_1 is a lexical verb	A has a schematic meaning, typically expressing distinctions of tense, aspect, or modality
	The subject may show semantic restrictions, such as being confined to animate participants	The subject shows no semantic restrictions
Morphosyntax	The subject is an argument of V_1	The subject is an argument of V_2
	V_1 belongs to the open class of lexical verbs	A belongs to a small paradigm of functional markers
	Being a lexical verb, V_1 is associated with the whole range of morphosyntactic options characteristic of verbs	Compared to the rich morphosyntactic potential of V_1, A disposes only of a severely limited range of morphosyntactic options
Phonology	Full form	A may be phonetically reduced vis-à-vis V_1

These observations are meant to show, first, that grammaticalization theory offers an account for understanding and describing the relationship between different types of grammatical constructions. With reference to the example used in this section this account can be summarized thus: Auxiliary constructions and the corresponding lexical constructions are related in a principled way via a set of parameters of grammaticalization.

Second, while this account can be phrased in terms of synchronic generalizations, it is based on diachronic regularities of grammatical change. Accordingly, the hypotheses proposed can be verified or falsified by means of diachronic evidence.

Third, these observations also suggest that there is no rigid boundary separating functional from lexical categories; rather, the transition between the two is notoriously fuzzy. As has been abundantly demonstrated in works on grammaticalization (see, for example, Heine and Kuteva 2002; 2006, ch. 2), lexical structures constantly give rise to new functional structures, and in a large number of cases it remains unclear whether a given construction qualifies as a lexical or a functional one. A survey of grammars of many languages shows that this has always been a problem that has vexed linguists; we will look at this issue in the next section.

To conclude, there is a systematic relationship between the non-grammaticalized construction and its grammaticalized product. Note, however, that this relationship is an asymmetrical one. While it is possible to derive the main properties of auxiliaries from their respective lexical sources, the opposite does not apply:[10] It is frequently not possible on the basis of the morphosyntactic properties of an auxiliary to reconstruct exactly the corresponding lexical verb.

For example, in hundreds of languages across the world, constructions having a verb of deictic motion "go to" as their predicate nucleus have been grammaticalized to a future tense auxiliary, the English *be going to*-future being a case in point. But there are equally many verbs for "come to" that have undergone the same grammaticalization process (see Bybee et al. 1994 for examples). On the basis of the morphosyntax of the resulting future tense categories it is as a rule not possible to infer whether the lexical source was a "go" or a "come" verb.[11]

16.6 Problems of grammatical taxonomy

The situation that we discussed in the preceding section raises the question of what the morphological or syntactic status of auxiliaries such as *keep* V-*ing, used to*, *have to*, or *be going to*, is or should be. Should they be treated as belonging to the same general taxon of lexical verbs or should they be assigned to a separate taxon, e.g., one of auxiliaries, or of functional categories? A survey of past linguistic analyses of the English language shows that this is an issue that has been discussed controversially. While some authors propose a main-verb account, others prefer to distinguish auxiliaries from lexical verbs (see Heine 1993: 8–9 for some positions that have been maintained on this issue).

Obviously, this is a theory-dependent issue, and which solution one favors is contingent upon the descriptive and/or explanatory framework one may wish to adopt. From the perspective of grammaticalization theory, neither the main-verb nor the autonomy hypothesis are entirely satisfactory; rather, we suggest that an auxiliary and the lexical verb from which it is historically derived jointly form a complex category whose structure is diachronically motivated. This category, having the format of what was referred to in section 16.4 as I/III structure, is characterized,

[10] It goes without saying that it is possible to reconstruct the lexical source of an auxiliary via its mophophonological form. Thus, the form of auxiliaries such as *keep* V-*ing* or *used to* allows us beyond any reasonable doubt to establish that it is the lexical verbs *keep* and *use*, respectively, that provided the lexical source of grammaticalization.

[11] An exception has been observed in the Bantu language Chaga of Tanzania, where both "go" and "come" have been grammaticalized to future tense categories and where these functional categories have retained distinguishing properties that bear witness to their respective pathways of grammaticalization (Emanatian 1992).

on the one hand, by a set of shared properties, in that the auxiliary retains some semantic, morphosyntactic, and morphophonological features of a lexical verb. On the other hand, it is characterized by a set of distinguishing properties which can be accounted for in a principled way via parameters of grammaticalization, such as desemanticization, decategorialization, and erosion (see section 16.3); accordingly, auxiliaries can be described roughly as desemanticized and decategorialized lexical verbs.

A problem of a different nature is raised by what we portrayed in section 16.4 as I/II/III structures. The issue here is not where a categorial boundary is to be traced but rather whether there is a boundary in the first place. We observed that reflexivity and reciprocity are different concepts, being assigned to different grammatical categories in many languages, including English, while in other languages, such as German, they are treated as belonging to the same morphosyntactic category but that within that category there is no boundary separating the two. That such a boundary problem can also be encountered across grammatical categories may be illustrated with the following example involving adpositions (that is, prepositions and postpositions).

There is a crosslinguistically common process of grammaticalization whereby noun phrases (or adverbial phrases) forming the head of a possessive/genitive construction develop into adpositions, and in some languages this is the primary source for adpositions. Underlying this grammaticalization there is a conceptual-semantic process whereby concrete, lexical concepts are pressed into service for the expression of more abstract, schematic functions of space, time, cause, purpose, etc. (see Heine 1997). In accordance with the overlap model in (6), this is a gradual process leading from stage I, where there is a nominal construction (A), via stage II, where this nominal construction can also be understood as an adpositional one (A/B), to stage III, where the construction is now exclusively an adpositional one (B). The result is that adpositions or, more commonly, a certain set of adpositions take the form of I/II/III structures—in other words, there is a grammaticalization chain extending from nominal-lexical to adpositional-functional structures.

As in the case of auxiliaries, the main features of adpositions can be described with reference to the parameters of grammaticalization. Desemanticization has the effect that the lexical meaning is bleached out, resulting in some schematic function. Decategorialization leads to the loss of categorial properties of the head noun, such as the ability to be inflected, to be pluralized, to take nominal determiners or modifiers, to be replaced by other nouns or a corresponding possessive pronoun, or to occur without their dependent noun phrase. And erosion may come in as well, leading to phonetic reduction of the adpositions.

We may illustrate the nature of I/II/III structures with a set of English complex (three-word) prepositions; Quirk et al. (1985) describe the situation of these prepositions thus:

Table 16.2. Relative degree of decategorialization of English complex prepositions based on nine indicators of syntactic separateness (based on Quirk et al. 1985: 671–2)

	Number of nominal properties	Examples
Maximally nominal	9	*on the surface of*
	4	*in defense of*
	3	*in comparison with*
	1	*in quest of*
Minimally nominal	0	*in spite of*

However, there is no absolute distinction between complex prepositions and constructions which can be varied, abbreviated, and extended according to the normal rules of syntax. Rather, there is a scale of "cohesiveness" running from a sequence which behaves in every way like a simple preposition [...] to one which behaves in every way like a set of grammatically separate units [...]. (Quirk et al. 1985: 671)

Via desemanticization, these prepositions lost their erstwhile lexical semantics, acquiring a schematic function of space (*in front of, in back of*), cause/reason (*on account of*), respect (*with reference to, with regard to*), instrument/means (*by means of*), purpose (*in order to*), concession (*in spite of*), comparison (*in comparison with*), etc. Decategorialization had the effect that the complex prepositions lost most of their nominal properties. As Table 16.2 shows, however, the prepositions differ in the extent to which they lost their nominal properties. At one end there is *in defense of*, which has retained many of the nominal properties distinguished by Quirk et al. (1985: 671–2), while *in spite of* is at the other end of the grammaticalization chain, having lost all its nominal properties.[12] We observed above that erosion is a parameter that may but need not be involved; it has also affected these English complex prepositions, in that they tend to be shortened in casual speech, cf. *(in) back of, (by) means of, (for) sake of, (in) spite of*, e.g.,

(14) Let's do this *sake of* consistency (Quirk et al. 1985: 671).

Note that these prepositions constitute but a range of the entire spectrum of grammaticalization. At a more advanced stage of the process, the three-word forms will be reduced to two-word forms and will become increasingly dissimilar to their lexical sources, as has happened with English *ahead of, because of, instead of*, etc. Arguably, this is already the case with *in spite of*; spelling conventions tend to be conservative and do not always reflect the actual degree of phonological and morphological fusion.

[12] Quirk et al. (1985: 671–2) describe decategorialization in terms of "indicators of syntactic separateness".

We were restricted here to one example relating to the interface between nominal and adpositional structures. Other examples are not hard to come by; suffice it to mention a similar kind of I/II/III structures linking verbal and adpositional structures. For example, a number of English verbs have been desemanticized in their gerundival form (*-ing*) and assumed prepositional functions, e.g., *barring, concerning, considering*, etc. Being desemanticized and having lost most of their verbal properties, such as to be inflected for tense and aspect, to take auxiliaries, etc., they in addition can be arranged along a scale extending from maximally verbal to prepositional forms (see Kortmann and König 1992).

16.7 Grammaticalization and crosslinguistic typology

In much the same way as overlapping categories of the type distinguished in section 16.4 pose problems for grammatical taxonomy based on "classical categories", they also concern linguistic typology, as the following example may show. Indefinite articles can be found in all parts of the world; according to a typological survey of 473 languages carried out by Dryer (2005*e*: 158–61; see also Dryer 1989), 43.1% of all languages have some kind of an indefinite marker. Dryer distinguishes the following types of articles (percentages relate to the number of languages in his sample of 473 languages):

(15) Typology of indefinite articles according to Dryer (2005*e*: 158)
a. Indefinite words that are distinct from the numeral "one" (19.2%),
b. numerals for "one" used as indefinite articles (19.0%),
c. indefinite affixes on nouns (4.9%).

As these observations suggest, indefinite articles exhibit a special relationship with the numeral "one", but Dryer does not address the question of how this relationship is to be explained. Research on grammaticalization has shown that there is an unambiguous answer to this question. In the vast majority of cases where languages have acquired an indefinite article this has happened via the grammaticalization of a numeral for "one". According to a worldwide survey of indefinite articles, using a genetically and areally balanced sample of 100 languages, Heine and Kuteva (2006: 99) found that 88.9% of all languages of the sample having an indefinite article have derived this article from a numeral for "one", and only 11.1% of the languages used other sources of grammaticalization.[13]

[13] But Heine and Kuteva (2006) found indefinite articles only in 18% of the sample languages, while in the remaining 82% of the languages there either was no conclusive information or it was

The development from numeral to indefinite article is in accordance with the parameters of grammaticalization. Desemanticization has the effect that the numeral meaning is bleached out, giving rise to a referential function (see Givón 1981); via decategorialization the numeral loses, for example, the ability to occur as a pronoun without a head noun, and erosion may come in as well, leading to a loss of segmental or suprasegmental phonetic substance (including the ability to be stressed). And this development is also in accordance with the overlap model in (6):[14] At stage I there is only a numeral (A), at stage II the numeral can be used as a marker for indefinite specific reference in specific contexts (A/B), and at stage III there is only an indefinite reference in some contexts (B).

European languages illustrate this process. Neither Classical Latin nor Proto-Germanic had an indefinite article; this was a stage I situation. In the course of the last two millennia, the descendants of these languages acquired articles in accordance with the overlap model—with the effect that in French or German, for example, there now are I/II/III structures (see (6) above): French *un/une* and German *einer/eine/ein* form overlapping categories characterized by many ambiguous (A/B) uses which can be interpreted simultaneously with reference to either the numeral or the article meaning. Thus, (16a) represents stage I of the overlap model, (16b) the ambiguous stage II, while (16c) is suggestive of stage III.[15]

(16) The overlap model and German *ein* "one, a"

a. Willst du ein Bier oder zwei?
want you one beer or two
'Do you want one beer or two?' (A)

b. Ich habe ein Fahrrad.
I have one/a bicycle
i 'I have one bike.' (A)
ii 'I have a bike.' (B)

c. Ein Bayer mag Bier.
a Bavarian likes beer
'Bavarians like beer.' (B)

explicitly stated that the languages made no use of indefinite articles. We have no conclusive answer on why there is such a discrepancy between the two samples, other than noting that Dryer (2005) apparently used a fairly liberal definition of indefinite articles. For example, he found an indefinite article in the Northern Khoisan language Ju-/'hoan, while our field research on this language in Namibia suggests that this language has neither a definite nor an indefinite article, and the same conclusion was reached by Dickens (1992; 2005: 21), who was Dryer's source of information.

[14] We are restricted here to the morphosyntactic development; the model applies equally well to the semantic-pragmatic development of the article (see Heine and Kuteva 2006).

[15] As in the case of the REF-REC category discussed in section 16.4, the grammaticalization from numeral to indefinite article involved loss of stress, that is, the parameter of erosion (see section 16.2). Accordingly, when *ein* receives stress, (16b) (and even (16c)) can only be interpreted in terms of the numeral meaning "one" of stage A.

But in other languages the process has gone one step further, especially with reference to erosion. This development was most dramatic in English, where the erstwhile numeral *one* was phonologically reduced to *a(n)*, thereby leading to a morphological split between numeral and article. The development was less dramatic in German and Dutch. In High German, erosion had the effect that the article lost the ability to be stressed, while in Dutch the full vowel of the numeral ([en]) was reduced in the article ([ən]).

With reference to the synchronic typology of (15), Dryer (2005: 158) allocates German and Dutch to different types on the following grounds: He argues that lack of stress is not reason enough to exclude German from type (b) of languages using the numeral as an indefinite article, while a reduced vowel is taken as a criterion to assign Dutch to type (a), that is, to languages where indefinite words are distinct from the numeral (note that in German the indefinite article is also commonly reduced phonetically in colloquial speech). Thus, on account of only a slightly different effect of erosion, German and Dutch are treated by Dryer as typologically unrelated, and Dutch appears in the same type as English.

But even in the case of English one may wonder whether it is entirely satisfactory to assign its indefinite article on the basis of a single morphophonological property to the same type (a) as languages having both etymologically and structurally contrasting words for the indefinite article and the numeral "one". On account of its history of grammaticalization, the English indefinite article shares a number of structural properties with the numeral "one": Like the numeral, it is restricted to singular head nouns (cf. *the trees* vs. **a trees*), it can be intensified by means of *single* (*in a single day*), it may occur in coordinate constructions (e.g., *a mile or two*), it exhibits a numerical or quantifying function in a number of other contexts (e.g., *a dozen, a quarter*), etc. (see Quirk et al. 1985: 273–4). Accordingly, the English indefinite article shares a number of properties with that of German and Dutch and at the same time contrasts structurally with type (a) languages where the numeral and the indefinite article are etymologically unrelated.

And much of what we observed on indefinite articles also applies to definite articles. In Dryer's (2005*d*: 154) typological analysis of definite articles there are 56 languages using a demonstrative word as marker of definiteness. We argue that these markers can be classified satisfactorily neither as demonstratives nor as definite articles but rather represent overlapping categories of the kind described above. What is required for a better understanding of their internal structure is an analysis of each of these markers in terms of the overlap model outlined in section 16.4.

To conclude, it would seem that both linguistic analysis and language typology might benefit from including findings on grammaticalization in order to arrive at a more comprehensive account of typological diversity among the languages of the world.

16.8 Conclusions

The preceding discussion has addressed a number of questions that are relevant to the general theme of the present volume. We saw that the main goal of grammaticalization theory is to understand why grammar in general and functional categories in particular are structured the way they are, and as in some other chapters of this volume (Givón; Bybee and Beckner) the conclusion reached is that this question cannot be answered satisfactorily without drawing on generalizations on language change in general and on grammaticalization in particular. Thus, typological generalizations on grammatical change provide a relevant tool for explaining structural properties of grammatical categories. Note, however, that the questions asked by students of grammaticalization differ from those raised in many other chapters of this volume, and so are the answers and explanations volunteered. Rather than aiming at a comprehensive theory of grammar, the approach sketched here is restricted to revealing how language is shaped by diachronic forces.

These forces are both communicative and cognitive in nature. Thus, one motivation of grammaticalization is to communicate successfully. Speakers draw on those discourse options that they consider most suitable for their communicative intents, use them more frequently, and they may exploit them for creating novel grammatical structures. To this end they draw on mechanisms of cognitive manipulation, such as metaphoric and metonymic transfer, using concepts of "concrete" domains of human experience to describe concepts of more abstract domains. For example, concepts of the domain of physical objects and actions are recruited to express non-physical concepts. Thus, one salient line of grammaticalization concerns the conceptual transfer from the domain of body parts to that of spatial organization. In a similar fashion, concepts for physical actions or motions, such as "do", "go", or "come", tend to be recruited for the expression of the more abstract concepts that are used to describe the temporal or aspectual contours of events, leading to the development of tense and aspect categories (Bybee et al. 1994); we drew attention to this process in section 16.5. The flip side of the cognitive aspects of grammaticalization is the social aspect, which it shares with other types of language change. Actual grammaticalizations are the result of innovations that were gradually adopted by a large number of speakers in a community. The spread of innovations in a community of speakers has been aptly described with the "invisible hand" metaphor by Keller (1994).

With regard to grammatical taxonomy, findings on grammaticalization suggest that traditional principles of classification in terms of "classical" categories based on necessary and sufficient conditions are not always well suited to account for overlapping structures such as the ones we discussed in the preceding sections. Rather than being exceptional or unusual, such structures are fairly common in any given language and in the languages of the world at large.

Grammaticalization concerns most parts of grammatical structure. Its core area is morphosyntax, that is, morphology and syntax and their interface, a second major interface area being that between the lexicon and grammar and between lexical and grammatical meanings, but for good reasons it has also been described as discourse manipulation having a pragmatic foundation. Phonology is mostly discussed in connection with issues of language processing and storage that pertain to grammaticalization (Bybee and Beckner, this volume; see also Bybee 2001*b*; 2002; 2003).

An area where grammaticalization has not been fully explored so far is that of sociolinguistics. Research in this field has concentrated on processes that are uniform across entire speech communities, and it has paid little attention to sociolinguistic variables. That processes of grammatical change are sensitive to social distinctions can be illustrated with an example that we were dealing with above (section 16.5): Speakers of English have created a range of complex prepositions in the course of the last centuries via the grammaticalization of adverbial phrases. This creativity appears to have been particularly pronounced in one social domain of language use, namely that of jurisdiction and bureaucracy, as is suggested by the fact that legal English stands out as having been particularly creative in the grammaticalization of new complex prepositions, such as *in case of, in default of, in lieu of, in pursuance of, in respect of*, or *on pain of* (Quirk et al. 1985: 672).

There is a clear answer to the question of what counts as evidence in the approach discussed in this chapter. While grammaticalization theory aims at accounting both for synchronic and diachronic phenomena, it has a diachronic foundation, in that it rests on generalizations of language use through time and space. Exemplification in this chapter concerned a few crosslinguistically regular diachronic processes, for example, from reflexive to reciprocal structures, from lexical verbs to auxiliaries, from nominal to adpositional structures, or from numeral to indefinite article, and the hypotheses we proposed can be verified or falsified by means of diachronic evidence. And these hypotheses relate both to typological diversity and to universal features of human languages. The processes discussed are based on typological comparisons and at the same time capture some general properties of human languages.

Another issue—one that was not dealt with in this chapter—concerns the question of whether or how the approach presented here deals with variation in language use and usage data. As has been demonstrated in a number of studies (see, for example, the contributions on English in Fischer (2000)), variation in language usage is a *sine qua non* for new grammatical structures to evolve, and some usage-based accounts have been proposed to deal with the process (see Bybee and Beckner, this volume; Bybee 2006*a*).

Recent research has established that grammaticalization is a ubiquitous phenomenon, to be observed in much the same way in spoken and written as in signed language use (Wilcox and Wilcox, this volume; Pfau and Steinbach 2005; 2006), and in language contact (Bisang 1996; Heine and Kuteva 2003; 2005; 2006). But there is one field that has so far received little attention in research on grammaticalization, namely that of language acquisition. On the basis of the observations made so far there is reason to assume that first language acquisition is in accordance with observations on grammaticalization (cf. Osawa 2003; Diessel 2005; Givón 2007; this volume), but more work is required on this issue.

This page intentionally left blank

CHAPTER 17

LEXICAL-FUNCTIONAL GRAMMAR

ASH ASUDEH
IDA TOIVONEN

17.1 Introduction

Lexical-Functional Grammar (LFG) was first developed in the 1970s by Joan Bresnan, a linguist at MIT, and Ron Kaplan, a psychologist at Harvard. Bresnan and Kaplan were concerned with the related issues of psychological plausibility and computational tractability. They wanted to create a theory that could form the basis of a realistic model for linguistic learnability and language processing. Since its foundation, the theory has been applied to numerous new areas, undergoing some modifications in the process, and has incorporated insights from a variety of morphological, syntactic, and semantic theories. However, the basic tenets of the theory and the formal framework have remained remarkably stable. For more on the history of LFG, see Kaplan (1987), Dalrymple et al. (1995: ix–5) and Dalrymple (2001: 1–5).

This work was supported by the Social Sciences and Humanities Research Council of Canada, Standard Research Grant 410-2006-1650. We would like to thank Mary Dalrymple, Bernd Heine, Daniela Isac, Helge Lødrup, Kumiko Murasugi, Heiko Narrog, Kenji Oda, Chris Potts, and Nigel Vincent for helpful comments and suggestions. We would also like to thank Daniela Isac and Cristina Moldoveanu for checking the Romanian data. Any remaining errors are our own.

LFG is a theory of generative grammar, in the sense of Chomsky (1957, 1965). The goal is to explain the native speaker's knowledge of language by specifying a grammar that models the speaker's knowledge explicitly and which is distinct from the computational mechanisms that constitute the language processor (Kaplan and Bresnan 1982). The central questions for LFG are thus largely the same as for other varieties of generative grammar: What is knowledge of language? How is it acquired? How is the knowledge embedded in a psycho-computational system? How do languages differ and how are they the same? The questions of acquisition and psychological processing were pursued particularly vigorously early in the theory's development; see various papers in Bresnan (1982*a*) and Pinker (1984). Computational questions have been investigated in detail in numerous publications, many of them stemming from work by the Natural Language Theory and Technology group at the Palo Alto Research Center (http://www2.parc.com/isl/groups/nltt/), as well as from work by research teams in Europe and Asia. The typological question of similarities and differences among languages has been particularly central to the subsequent development of the theory.

In answering these questions, LFG research draws on a wide variety of evidence: native speaker intuitions, corpora, psycholinguistic evidence, typological patterns, and computational models. The Xerox Linguistic Environment (XLE; Crouch et al. 2008) is a robust computational implementation of LFG that has allowed explicit testing of theoretical hypotheses, leading to new research areas and formal innovations in the process. The development of XLE has led to computational work on efficient parsing (e.g., Maxwell and Kaplan 1991, 1993, 1996). XLE also forms the basis for a variety of industrial applications, such as the Powerset search engine, which is based on linguistically sophisticated natural language understanding (as opposed to the more superficial "bag of words" approach that is the norm).

A central idea of Lexical-Functional Grammar is that different kinds of linguistic information are modeled by distinct, simultaneously present grammatical modules, each having its own formal representation. The grammatical architecture of LFG thus postulates a number of simple data structures with mappings defining the relationships between structures. The different grammatical modules are subject to separate principles and formal descriptions and have distinct primitives. However, at the heart of the architecture are simple set-theoretic concepts. The structures are defined in terms of sets of primitive elements and functions and relations on these sets. The mappings between structures are also defined in terms of functions and relations. LFG's formal architecture is thus typically referred to as a *Parallel Projection Architecture* or *Correspondence Architecture* (Kaplan 1987, Halvorsen and Kaplan 1988, Kaplan 1989, Asudeh 2006),[1] because different grammatical

[1] In order to avoid potential confusion with the distinct "Parallel Architecture" developed by Jackendoff (1997, 2002), we will use the latter name.

components are present in parallel and correspond to or are projected to each other by what are alternatively called correspondence or projection functions. This kind of architecture contrasts strongly with architectures in which different kinds of grammatical information are modeled by identical data structures and are subject to the same operations. LFG can be contrasted, for example, with some versions of Principles and Parameters Theory, where morphological, syntactic, and semantic information alike are modeled with phrase structure trees and where phrases, words, morphemes, and features alike are combined with the same operations for insertion and manipulation of syntactic structures. Similarly, in Head-Driven Phrase Structure Grammar, all linguistic information is modeled by directed acyclic graphs (Pollard and Sag 1994). LFG is often viewed as a syntactic framework, but it is important to note that other modules of grammar have also been developed within the correspondence architecture of LFG. We return to this topic in section 17.5.

17.2 C-STRUCTURE AND F-STRUCTURE

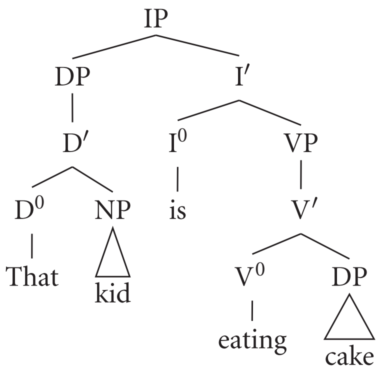
LFG posits two syntactic structures: constituent structure (c-structure; occasionally also referred to as *categorial structure*) and functional structure (f-structure). This section describes the two structures and presents the linguistic intuitions that lie behind their separation.

C-structures are represented as phrase structure trees and model precedence (word order), dominance, constituency, and syntactic categories. F-structures are represented as feature structures (also known as attribute value matrices). An f-structure is a finite set of attribute–value pairs, such that an attribute is a symbol and its value is: (a) a symbol (e.g., SINGULAR or +); (b) a semantic form (a potentially complex symbol in single quotes); (c) a set; or (d) an f-structure. The f-structure of a sentence contains the grammatical functions that the head verb subcategorizes for (SUBJECT, OBJECT, etc.) and also represents a range of morphosyntactic information, such as case, agreement features, tense, and aspect. F-structure is the level at which abstract syntactic relations are captured, such as agreement, control and raising, binding, and unbounded dependencies.

Turning to an example, the c-structure and f-structure for (1) are shown in (2) and (3) respectively.

(1) That kid is eating cake.

(2)



(3)
$$\begin{bmatrix} \textsc{pred} & \text{‘eat}\langle\textsc{subj}, \textsc{obj}\rangle\text{’} \\ \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{‘kid’} \\ \textsc{deixis} & \textsc{distal} \\ \textsc{definite} & + \\ \textsc{number} & \textsc{singular} \\ \textsc{person} & 3 \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{‘cake’} \\ \textsc{number} & \textsc{singular} \\ \textsc{person} & 3 \end{bmatrix} \\ \textsc{tense} & \textsc{present} \\ \textsc{aspect} & \textsc{progressive} \\ \textsc{participle} & \textsc{present} \end{bmatrix}$$

The vertical order of features in an f-structure is not important, since an f-structure is just an unordered set of attribute–value pairs. A richer example f-structure can be found in appendix A.

F-structures can be compared to the "relational networks" of Relational Grammar (Perlmutter and Postal 1977, Perlmutter 1983), since both structures model grammatical functions (or relations). However, the formalization is very different. First, the values of LFG grammatical functions are feature structures which contain morphosyntactic information, such as case and agreement features. In general, LFG f-structures thus tend to contain considerably more grammatical information than relational networks. Second, relational networks represent relation changes in tiered strata—where subsequent strata are derived derivationally from previous strata. In contrast, such changes are not represented in f-structure, since a key tenet of LFG theory is that all relation changes are lexical (Bresnan 1978, 1982*b*, Kaplan and Bresnan 1982; see Bresnan (2001*b*: 25–40) for an overview of some of the main arguments).[2]

[2] Role and Reference Grammar (RRG; Van Valin 1993b, 2005) also posits grammatical functions, as syntactic arguments tied to semantic roles such as Actor and Undergoer. RRG is based on quite

C-structure models the surface exponence of syntactic information, such as word order and constituency, whereas f-structure models more abstract syntactic information and relations. C-structures may vary widely between languages, but f-structural information remains relatively constant across languages. It is thus at f-structure that we observe many cross-linguistic universals. Consider passives, for example. Perlmutter and Postal (1977) show that it is not possible to describe the passive cross-linguistically with reference to verbal morphology, case marking, or word order. What regular passives[3] have in common cross-linguistically is that the subject is demoted and the object is promoted to subject. However, not all languages mark their subjects and objects the same: in some languages, subjects are distinguished from other functions with case marking, in some with agreement on the verb, and yet others distinguish between these (and other) grammatical functions with word order and phrase structure. Of course, many languages use a combination of several linguistic devices to distinguish between grammatical functions. F-structure directly models grammatical functions, such as subjects and objects, whereas c-structure displays the more superficial information about how the functions are encoded in a given language. The LFG analysis of passives and other relation changes is captured in the mapping between argument roles (such as *agent*, *patient*, etc.) and grammatical functions. The difference between an active and a passive sentence lies in which argument is realized as the subject at f-structure. How this subject is expressed in c-structure is language-specific. The theory of these mappings was initially developed in the 1980s and 1990s and is called Lexical Mapping Theory (Levin 1985, Bresnan and Kanerva 1989, Alsina 1996, Butt 1995, Butt and King 2000*a*), and this has been a very active area of research in LFG. See Bresnan (2001*b*: 318–20) for numerous additional references. The remainder of this section examines some characteristics of c-structure and f-structure in turn.

17.2.1 C-structure

C-structures are constrained by the principle of Lexical Integrity (see Bresnan (2001*b*: 91–3) for an overview):

(4) **Lexical Integrity**
The terminal nodes of c-structures are morphologically complete words.

This has two immediate consequences. First, terminal nodes in c-structure cannot be morphemes or morphological structures smaller than words, in

different conceptual foundations from LFG, since the former is a functionalist linguistic theory and LFG is not. See Farrell (2005) for a comparison of grammatical relations in LFG and RRG.

[3] The "regular" passive can be compared to *impersonal passives*, where the object is not promoted, and *pseudo-passives*, where a prepositional object is promoted.

contrast to what obtains in certain other theories (e.g., Distributed Morphology; Halle and Marantz 1993, Embick and Noyer 2007, among others). The syntax is therefore blind to the internal structure of words and sees only their category. This has a number of further consequences, which are explored in the LFG literature. For example, morphological units are correctly predicted not to support certain syntactic operations, such as extraction, gapping, coordination, certain anaphoric dependencies, and certain kinds of recursion. These consequences of Lexical Integrity are considered by Bresnan and Mchombo (1995), who show that Lexical Integrity provides a principled explanation of the complex syntactic, morphological, and prosodic properties of Bantu noun class markers.

A second consequence of Lexical Integrity, which has not thus far received as much attention in the LFG literature (although, see Asudeh et al. 2008), is that terminal nodes in c-structure cannot be syntactic units larger than morphological words. That is, the lexicon does not provide c-structure with fully formed phrases; compare, for example, the lexically stored phrases of Construction Grammar (Fillmore 1988, Goldberg 1995, Kay and Fillmore 1999).

Pre-terminals are labeled with the syntactic category of the word that fills the terminal node. The set of category labels includes a number of lexical categories: N(oun), V(erb), P(reposition), A(djective), and Adv(erb). Many LFG researchers have also adopted a limited set of functional categories and projections (see, for example, Kroeger 1993, King 1995, Bresnan 2001*b*, Dalrymple 2001). The functional categories assumed are typically C(omplementizer), I(nflection), and D(eterminer). In general, the only functional categories adopted in LFG are ones involved in word order and distributional generalizations. For example, the categories C and I are involved in LFG analyses of head displacement phenomena, such as verb-second in Germanic languages, and the distribution of English auxiliaries. Functional categories such as K (Case) and Agr(eement) are therefore not adopted,[4] since information about case and agreement is captured in the morphology and at f-structure.

The *exocentric* (i.e., lacking a phrase structure head) category S is widely adopted within LFG. It serves two purposes. First, it is used in analyses of languages that lack a VP and display a flat constituent structure, such as Warlpiri (Simpson 1983, 1991, Nordlinger 1998, Bresnan 2001*b*). Second, it is used in analyses of [$_S$ NP XP] predication structures, where the predicate phrase XP may be VP, NP, AP or PP. These sorts of predication structures are common in Celtic languages (see, for example, Chung and McCloskey 1987 for Irish).

[4] There are exceptions. For example, Butt and King (2004) adopt the category K for Hindi–Urdu case endings, because they argue that these are clitics and they want to maintain a generalization that only functional heads may be clitics.

17.2.2 F-structure

One of the principal motivations for the name *functional* structure is the fact that grammatical functions are represented at f-structure. A second motivation is that functional structures are finite functions in the mathematical sense, due to the condition on f-structure wellformedness known as Uniqueness or Consistency:[5]

(5) **Uniqueness/Consistency**
Every f-structure is such that every attribute has exactly one value.

F-structures are thus total functions from attributes to values. However, they may be many-to-one functions: different attributes may have the same value. Shared values can be observed in standard LFG analyses of raising, obligatory control, and unbounded dependencies.

Grammatical functions are a reflection of predicate–argument relations, and a central purpose of f-structure is to capture these relations. One motivation for this is the typological observation that nonconfigurational languages (e.g., Warlpiri) encode similar predicate–argument relations to configurational languages. A nonconfigurational language and a configurational language may have the same f-structure corresponding to strikingly different c-structures; see Bresnan (2001*b*: 5–10) for an expository discussion of this point with respect to Warlpiri and English. A second, overarching motivation is the observation that many syntactic phenomena can be compellingly analyzed in terms of predicate–argument relations (cf. the discussion of passives above). A distinguishing feature of LFG is its adoption of a rich inventory of grammatical functions as primitives of the theory. Table 17.1 contains an overview of LFG's grammatical functions.

The grammatical functions (GFs) in Table 17.1 can be cross-classified in a number of ways. First, a subset of the grammatical functions—the *governable grammatical functions*—may be directly selected by predicates.

(6) **Governable grammatical functions:** SUBJ, OBJ, OBJ_θ, OBL_θ, COMP, XCOMP

All other GFs are non-governable and cannot be specifically selected, but rather occur freely, subject to other constraints of the theory.

The core nominal grammatical functions are further decomposed in Lexical Mapping Theory according to the features [± r(estricted)] and [± o(bjective)], as follows:

(7) **LMT decomposition of core nominal GFs**

	−o	+o
−r	SUBJ	OBJ
+r	OBL_θ	OBJ_θ

[5] LFG is not committed to "functional" in the sense of functionalist linguistics, namely having to do with communicative functions of language.

Table 17.1. Grammatical functions in LFG

SUBJECT (SUBJ)	Some people with no shame walked in and wrecked the party. The party was wrecked by some people with no shame.
OBJECT (OBJ)	*First object.* Ricky trashed the hotel room. Ricky gave John a glass. Ricky gave a glass to John.
OBJECT$_\theta$ (OBJ$_\theta$)	*Second object. Thematically restricted object.* Sandy gave John a glass. Tom baked Susan a cake. #Tom baked a cake Susan. (OBJ$_\theta$ *in English restricted to theme, cannot be beneficiary*)
OBLIQUE$_\theta$ (OBL$_\theta$)	*A complement (non-subject argument) that has oblique case or is a PP.* Julia placed the vase on the desk. Ricky gave a glass to John.
COMPLEMENT (COMP)	*Closed (saturated) complement: A clausal argument which has its own subject.* Peggy told Matt that she had won the prize.
XCOMP	*Open (unsaturated) predicate complement: A predicative argument with no overt subject of predication.* I told Patrick to quit. Peggy-Sue seems to be a complete fraud.
ADJUNCT (ADJ)	*A modifier, a non-argument.* Mary read a good book. Mary counted the cars very quickly. Sally killed a bug in the yard. Since she had no money, Mary was forced to get a job.
XADJ	*Open predicate adjunct.* Having no money, Mary was forced to get a job.
SPECIFIER (SPEC)	*Possessor or quantificational determiner phrase.* John's book's cover is red. At least three books are red.
TOPIC (TOP)	*Grammaticalized discourse function.* *Must be identified with or anaphorically linked to another grammatical function.* Mary met the author whose books ___ annoyed Peggy. (TOPIC=SUBJ) Bagels, Mary loves ___. (TOPIC=OBJ) As for bagels, Mary loves them. (TOPIC anaphorically linked to OBJ)
FOCUS (FOC)	*Grammaticalized discourse function.* *Must be identified with or anaphorically linked to another grammatical function.* Which author do the critics praise ___? (FOCUS=OBJ) Cén t-údar a molann na léirmheastóirí é? Which author COMP praise the critics him (FOC anaphorically linked to OBJ) (Irish; McCloskey 1979: 53)

Grammatical functions that are tied to specific thematic roles are [+r], whereas functions that are semantically unrestricted are [−r]. For example, subjects and direct objects can bear any semantic role, whereas second objects are typically restricted to patients and themes. Subjects and objects are also unrestricted in another sense: they can be expletives, lacking a semantic role entirely. The feature [+o] distinguishes the object functions, OBJ and OBJ$_\theta$, from the subjects and obliques.

A second cross-classification of grammatical functions is according to whether the GF is *closed* or *open*. The open grammatical functions are the open complement function XCOMP and the open adjunct function XADJ. Open grammatical functions contain a predicate that requires a subject of predication, but which depends on external specification of the subject through the functional predication relation known as *functional control*. A functional control equation relates the XCOMP or XADJ's subject to a grammatical function in the f-structure in which the XCOMP or XADJ occurs (Bresnan 1982*c*). A typical instance of an XCOMP is the complement of a raising verb:

(8) Alfie seemed to laugh.

(9) Alfie vaikutti nauravan. Finnish
Alfie seemed laugh
Alfie seemed to laugh.

(10)
$$\left[\begin{array}{ll} \text{PRED} & \text{‘seem}\langle\text{XCOMP}\rangle\,\text{SUBJ’} \\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{‘Alfie’}\end{array}\right] \\ \text{XCOMP} & \left[\begin{array}{ll}\text{PRED} & \text{‘laugh}\,\langle\text{SUBJ}\rangle\text{’} \\ \text{SUBJ} & \end{array}\right]\end{array}\right]$$

The connecting line in the f-structure represents the functional control relation, which in this case is lexically specified by the raising verb *seemed* (informally: “my SUBJ is my XCOMP's SUBJ”). In contrast, the functional control relation for an English XADJ would be associated with a c-structure position rather than specified lexically, since XADJ is a non-governable grammatical function that is not selected for, but rather appears freely, like other adjuncts, in certain structural positions.

A third cross-classification of grammatical functions is according to whether they are grammaticalized discourse functions or not. The discourse functions are typically structurally prominent in some way, as an expression of their discourse prominence.

(11) **Discourse functions:** TOPIC, FOCUS, SUBJ

Subject is the only discourse function that is also a governable GF. Topic and FOCUS are not selected directly but are rather integrated into the f-structure by the

Extended Coherence Condition (see (55) below). In many languages, the SUBJ is also identified by default as the TOPIC. See Falk (2006) for a recent in-depth LFG-theoretic investigation of subjects.

Grammatical functions are subcategorized for in lexically specified PRED features, which we have already encountered in (3) and (10). For example, the verb *eating* has the PRED value 'eat⟨SUBJ,OBJ⟩'. The first part of this value is the predicate function, which is conventionally the stem form of the lexical item that contributes the PRED. It is also a common convention for the predicate function to be written in a convenient meta-language for the linguist, rather than in the language of analysis. For example, the PRED value for the Finnish *vaikutti* in (9) is 'seem…'. The PRED feature also specifies how many and which governable grammatical functions the verb selects, as indicated in its argument list (the grammatical functions specified after the predicate function).

Lastly, a distinction is drawn between thematic and non-thematic arguments. Thematic arguments are written within the angled brackets, whereas non-thematic arguments are written following the angled brackets. For example, the PRED of *seemed* in (10) is 'seem⟨XCOMP⟩SUBJ', which indicates that the XCOMP complement is the only thematic argument of *seem* and that the raising verb's SUBJ is non-thematic. There is a general requirement that thematic arguments must have a PRED feature, since they are semantically contentful, whereas non-thematic arguments need not have a PRED, since they are not semantically contentful. For example, the SUBJ of a raising-to-subject verb may be an expletive. Expletives in LFG are analyzed as lacking a PRED feature but having appropriate agreement features. It is important to realize, though, that PRED is not a semantic representation but rather the syntactic exponent of certain semantic information.

The value of a PRED attribute is a *semantic form*, which is indicated by the enclosing single quotes. Semantic forms are special, complex symbols that are always uniquely instantiated. This is captured formally through indexation on semantic forms, e.g. 'eat$_{12}$⟨SUBJ,OBJ⟩', but the indices are typically suppressed. Unique instantiation of semantic forms ensures that semantically relevant information is not multiply specified syntactically. For example, consider the following examples from Irish (McCloskey and Hale 1984: 489–90):

(12) Chuirfinn isteach ar an phost sin. — Irish
put.COND.1SG is on that job
I would apply for that job.

(13) *Chuirfinn mé isteach ar an phost sin.
put.COND.1SG I is on that job

Irish has both *synthetic* and *analytic* forms of verbs in certain paradigms. Synthetic forms contain complete pronominal information and cannot occur with an overt pronominal, even if the pronoun is compatible in agreement features with the verb. *Chuirfinn* is the synthetic form of the conditional form of *cuir* ("put") in the first person singular. Example (13) is thus ungrammatical because the synthetic verb form cannot occur with the overt pronominal. Andrews (1990) shows that this falls out neatly from the uniqueness of semantic forms in LFG. The synthetic verb form incorporates pronominal information (McCloskey and Hale 1984) and therefore contributes its SUBJECT's PRED feature, specifying its value as 'pro' (the standard LFG-theoretic PRED value for non-expletive pronominals). The independent pronoun *mé* also contributes a PRED feature with value 'pro'. However, the two instances of 'pro' are unique semantic forms and thus cannot simultaneously be the value of a single PRED feature. This results in a violation of Consistency, defined in (5) above. Example (13) is thus correctly predicted to be ungrammatical. The situation exemplified here can be contrasted with "pro-drop" languages, in which the verb's contribution of its SUBJ's PRED is optional and the verb therefore may appear with a suitably agreeing overt subject; see the Romanian examples in (42–44) below. The theory thus derives the distinction between obligatory suppression of a pronominal subject, as in these Irish cases, from optional suppression of a pronominal subject, as in Romanian, based on obligatoriness vs. optionality of relevant lexical information.

In addition to Consistency, there are two other general wellformedness conditions which apply to all f-structures:

(14) **Completeness**
An f-structure is *complete* if and only if it contains all the governable grammatical functions that its predicate governs.

(15) **Coherence**
An f-structure is *coherent* if and only if all the governable grammatical functions it contains are governed by a predicate.

Note that the term "govern" means nothing more than to be listed in the argument list of a PRED feature.

Completeness and Coherence serve a similar role in LFG as the Projection Principle, the Theta Criterion, and Full Interpretation do in P&P and that the Subcategorization or Valence Principle does in HPSG. They ensure that the subcategorization requirements of a predicate are met exactly. Coherence violations occur if a constituent cannot be mapped to any GF (i.e., if there are "extra" arguments):

(16) *Thora remembered every movie most videos.

(17) *That the earth is round did not surprise Columbus that he could sail west without danger.

Completeness violations occur if subcategorized GFs are not present, as in the following examples:

(18) *Alfie devoured.

(19) *John wondered if seemed to be a problem.

Example (19) illustrates that Completeness requires even non-thematic governed GFs to be present. Even though the SUBJ of *seemed* is non-thematic it is still required by Completeness; that is, Completeness applies to *all* GFs in a PRED's argument list, both inside and outside the angled brackets.

17.3 Structures and structural descriptions

LFG distinguishes sharply between formal structures, such as c-structures and f-structures, and structural descriptions that wellformed structures must satisfy. The structural descriptions are sets of constraints. A constraint is a statement that is either true or false of a structure. This section provides an overview of the most important sorts of constraints. For a more thorough discussion, see in particular Dalrymple (2001: 91–176).

17.3.1 Constraints on C-structures

The formal structures in c-structure are phrase structure trees, as illustrated in (2) above. The structural descriptions that constrain the phrase structure trees are formalized as phrase structure rules, such as (20):

(20) $\text{IP} \rightarrow \text{DP} \quad \text{I}'$

A wellformed c-structure must satisfy all applicable phrase structure rules and every sub-tree in a wellformed c-structure must satisfy some phrase structure rule. The body of LFG's phrase structure rules are *regular expressions*, which support optionality, disjunction, negation, and arbitrary repetition. Regular expression repetition uses the Kleene operators (Kleene 1956): Kleene star (*), which means "zero or more occurrences of the annotated expression", and Kleene plus ($^+$), which means "one or more occurrences of the annotated expression". LFG's phrase structure rules are comparable, in this specific respect, to the phrase structure rules of Generalized Phrase Structure Grammar (GPSG; Gazdar et al. 1985), which also support regular expressions (Gazdar et al. 1985: 54–5). A formal exposition of regular expressions can be found in Partee et al. (1993: 462–4).

Consider, for example, the following V′ rule, proposed solely for illustration:

(21) $\text{V}' \rightarrow \text{V}^0 \quad (\text{NP}) \quad (\{\text{CP}|\text{VP}\}) \quad \text{PP}^*$

Optionality is indicated by parentheses around a rule element. Disjunction is indicated with the notation {X|Y}. Rule (21) has a single obligatory element, the V^0.

The verb may be immediately followed by an NP sister. The verb may also have either a CP or a VP sister or neither (since the entire disjunction is within the scope of optionality parentheses). Lastly, the V′ may end in any number of PPs, including none.

Phrase structure rules are posited separately for independent languages, subject to certain universal principles. A structure is allowed only if it is linguistically motivated for that language. The motivation consists primarily of distributional evidence (for category assignment), constituency evidence, and word order. For example, if the verb appears after its complements in a given language, the VP rule for that language is V-final. There is no attempt to derive all surface word orders from a universal underlying word order, such as SVO (Kayne 1994); this notion makes no sense in LFG, since the theory is not derivational and does not postulate underlying word order that is distinct from surface word order. LFG's "surface-true" approach to phrase structure is further evidenced by the fact that a VP is posited only if there is distributional or constituency evidence for such a category. A language without a VP is a non-configurational language (see Nordlinger 1998 and references cited therein for definitions of non-configurationality).

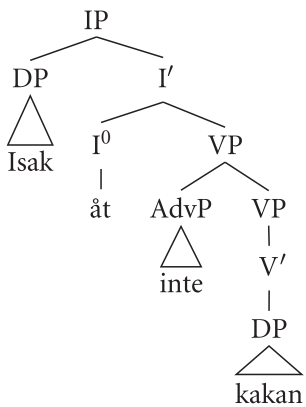
Although c-structures vary greatly cross-linguistically, the variation seems to be limited in a principled way. This is captured in LFG with X-bar theory (Chomsky 1970, Bresnan 1977, Jackendoff 1977) and certain universal principles on the c-structure to f-structure mapping. The mapping principles are discussed in detail in Bresnan (2001*b*: 98–109) and Toivonen (2003: 66–9). One principle states that "c-structure heads are f-structure heads". This means that a c-structure head maps its featural information into the same f-structure as its c-structural mother. Such principles sharply limit the combinatorial possibilities at c-structure.

LFG allows for both *endocentric* c-structures and *lexocentric* c-structures, the latter rooted in the exocentric category S, as discussed in section 17.2.1. Lexocentric phrase structure is instantiated in languages where grammatical functions are encoded morphologically rather than configurationally. Lexocentric structure is both typologically common and diverse (instantiated in genetically and geographically unrelated languages). However, the theory assumes that S is the only exocentric category and that, even within lexocentric languages, other categories are endocentric. The theory thus posits S as a principled exception to X-bar theory in order to capture phrase-structural properties of lexocentric languages without forcing them into a configurational mold.

LFG's use of X-bar theory provides a good illustration of the conceptual difference between structures and structural descriptions. Consider a typical LFG analysis of "head movement" phenomena (Travis 1984), which in LFG do not involve movement at all but rather lexical specification of a functional category such as I^0 for a verb (King 1995). For example, consider Germanic verb-second, as instantiated in Swedish:

(22) Isak åt inte kakan. Swedish
Isak ate not cookie.DEF
Isak did not eat the cookie.

The Swedish finite verb in this example has the category I^0 in the lexicon and is thus analyzed as base-generated in I^0 (Sells 2001*a*, Toivonen 2001, 2003), yielding the following structure:

(23)


The V′ in (23) does not contain a V^0, a violation of X-bar theory as a theory of c-structures. However, the relevant phrase structure rule—which a tree rooted in V′ must satisfy—does contain a V^0, although an optional one:

(24) $V' \rightarrow (V^0) \quad \ldots$

Thus, X-bar theory in LFG holds as a theory of structural descriptions. For more detailed discussions of X-bar theory and LFG's theory of phrase structure, see Bresnan (2001*b*: chapter 6) and Toivonen (2003: chapter 3).

Lastly, LFG's theory of c-structure does not posit any principle that dictates that multiply branching structures are disallowed. For example, both objects of a ditransitive verb are sisters of the verb. Coordination structures are also multiply branching. LFG rejects the contention that all phrase structure is binary branching (Kayne 1984). The putative evidence for that claim concerns phenomena that are analyzed at f-structure.

17.3.2 Constraints on F-structures

F-structure constraints are stated in a quantifier-free theory of equality. F-structure constraints are specified in lexical entries and in annotations on nodes of c-structures, as explained in more detail in section 17.4 below. The set of all f-structure constraints obtained from the lexical entries and c-structure of a given analysis is called a *functional description* or *f-description*.

A common kind of constraint is a *defining equation*, which specifies the value of some attribute in an f-structure. For example, the following defining equation specifies that the NUMBER attribute of some f-structure f has the value SINGULAR:

(25) $(f\ \textsc{number}) = \textsc{singular}$ *defining equation*

The values of f-structures can also be semantic forms and other f-structures, so we also get these sorts of defining equations:

(26) $(f\ \textsc{pred}) = \text{‘laugh}\langle\textsc{subj}\rangle\text{’}$

(27) $(f\ \textsc{subj}) = g$

The equation in (27) states that the SUBJECT of f-structure f is f-structure g.

Recall that f-structures are functions. Thus, an equation such as (25) can be understood as a kind of functional application, where we write the parentheses as above instead of the more standard (28):

(28) $f(\textsc{number}) = \textsc{singular}$

The reason this notional difference was instituted is that it makes iterative functional applications easier to understand. For example, consider the partial f-structure, f, in (29).[6]

(29)
$$f: \begin{bmatrix} \textsc{pred} & \text{‘smile}\langle\textsc{subj}\rangle\text{’} \\ \textsc{tense} & \text{present} \\ \textsc{subj} & g\begin{bmatrix} \textsc{person} & 3 \\ \textsc{number} & \textsc{singular} \end{bmatrix} \end{bmatrix}$$

Now suppose that (29) represents part of the f-structural information of sentence (30) and we want to specify subject–verb agreement.

(30) Alfie smiles.

We can capture the agreement by specifying the following two equations in the verb's lexical entry:

(31) $(f\ \textsc{subj number}) = \textsc{singular}$
$(f\ \textsc{subj person}) = 3$

Given that f's SUBJ is g in (29), these simplify to:

(32) $(g\ \textsc{number}) = \textsc{singular}$
$(g\ \textsc{person}) = 3$

[6] We have written a colon after the f-structure label f to make clear that the f-structure is the function f. We will henceforth suppress the colon, but a label f on an f-structure should be read as the name of the f-structure, not as a function f applied to an unnamed f-structure.

These equations will only be satisfied if the subject *Alfie* is compatible with the number and person specifications indicated in (32). Since this is indeed the case, (30) is correctly predicted to be grammatical. In sum, successive functional applications can be represented by writing what amount to *paths* of attributes.

Defining equations can be expressed in terms of relations other than basic equality. One common relation is set membership, since at f-structure modification is represented as a set that is the value of an ADJUNCT grammatical function. ADJUNCTS are represented as sets because there is no upper bound on the number of modifiers that a constituent may have.

(33) $g \in (f \text{ ADJ})$

This equation states that the f-structure g is a member of the ADJ set of f-structure f; see Appendix A for an f-structure containing ADJUNCT sets. Sets are also used in the f-structural representation of coordination (Kaplan and Maxwell 1988, Maxwell and Manning 1996) and in a more articulated theory of morphosyntactic features that accommodates resolution of coordinated morphosyntactic information (Dalrymple and Kaplan 2000).

The solution for a given f-description is the *minimal* f-structure that satisfies the set of constraints. The minimal f-structure contains only the attributes and values that are explicitly mentioned in the f-description. If the minimality constraint did not hold, then the f-structure for (34) would equally satisfy the f-description for (35), since the additional modifier *quickly* contributes information that is not inconsistent with the smaller f-structure.

(34) Alfie quickly ran out.

(35) Alfie ran out.

However, it is clear that we would not want (34) and (35) to have the same f-structural parse, because they are syntactically distinct sentences.

A second kind of equation, the *constraining equation*, takes advantage of the minimality requirement. Constraining equations do not define the features and relations in an f-structure but rather check that the minimal f-structure has the features or relations specified by the constraining equation. Formally, the constraining equations are evaluated once the set of defining equations has been satisfied by the minimal f-structure. A constraining equation is written with a subscripted c:

(36) $(f \text{ PARTICIPLE}) =_c \text{ PRESENT}$ *constraining equation*

This equation does not result in f-structure f having the feature PARTICIPLE with value PRESENT. Rather, it checks that f contains that feature and value. An independent defining equation must actually specify the feature and value.

In order to see how this is useful, consider these examples (following a similar discussion in Kaplan and Bresnan 1982):

(37) Thora is giving Harry a toy.

(38) *Thora is gives Harry a toy.

Let us assume that the progressive auxiliary *is* and the participle *giving* map to the same f-structure, f, and that the constraining equation (36) is part of the lexical information associated with the auxiliary. Let us make the natural assumption that the present participle *giving* has a PARTICIPLE feature with value PRESENT, lexically specified through a defining equation associated with the participle. See (65) in section 17.4.1 below for the relevant f-descriptions. Since the auxiliary's constraining equation is thus satisfied in (37), the sentence is correctly predicted to be grammatical. In contrast, let us assume that the present tense form *gives* does not specify any participial information, since it is not a participle form of the verb. Example (38) is thus ruled out, because the auxiliary's constraining equation cannot be satisfied, since *gives* does not provide the required information.

Now consider what would be the case if (36) were a defining equation rather than a constraining equation. If *gives* did not provide any information to the contrary, then the progressive auxiliary would actually just add the feature PARTICIPLE with value PRESENT and (38) would incorrectly be predicted to be grammatical. In order to block (38), every non-participial verb would have to specify a participle feature with a value such as NIL or NON-PARTICIPIAL. The constraining equation allows us to avoid this inelegant and unnatural situation, since only participles need be marked as such. This participial example demonstrates one of the key uses of constraining equations, which is to control co-occurrence of words or phrases through their associated f-structure features.

There are three other useful kinds of constraints on minimal solutions. *Negative equations* are satisfied if and only if a feature has a value other than the one specified (including complete absence of the feature):

(39) $(f\ \text{CASE}) \neq \text{NOMINATIVE}$ or $\neg[(f\ \text{CASE}) = \text{NOMINATIVE}]$ *negative equation*

The first notation is somewhat more common. The negative equation (39) is satisfied if and only if f has no CASE feature or if the value of CASE is something other than NOMINATIVE.

The last two kinds of constraint are the *existential constraint*, which is satisfied if and only if the attribute in question is present (regardless of its value), and the related *negative existential constraint*, which is satisfied if and only if the attribute in question is absent (regardless of its value). Here is an example of each kind of constraint:

(40) $(f\ \text{CASE})$ *existential constraint*

(41) $\neg(f\ \text{CASE})$ *negative existential constraint*

The existential constraint (40) requires f to have a CASE feature. The negative existential constraint (41) requires f not to have a CASE feature.

The Boolean connectives of conjunction, disjunction, and negation can be used in f-descriptions. Conjunction is typically implicit: in any f-description, all the constraints must hold. Conjunction can also be explicitly indicated with the standard symbols "&" or "$\wedge$". Disjunction is indicated either with the symbol "$\vee$" or in the form "{X | Y}". Negation is indicated with the symbol "$\neg$". Grouping is indicated by square brackets, "[...]". Optionality is once again indicated by parentheses, "(...)".

Judicious use of these connectives allows for compact specification of f-structure constraints. For example, consider the following two examples from Romanian, a pro-drop language that shows syncretism of first and second person singular in certain conjugations (Cojocaru 2003: 120–6):

(42) Eu/tu continui. Romanian
I/you continue.PRES.[1.SG/2.SG]
I/you continue.

(43) Continui.
continue.PRES.[1.SG/2.SG]
I/you continue.

(44) *Ea continui.
she continue.PRES.[1.SG/2.SG]

The verb *continui* ("continue") lexically contributes the following f-description, where f is the f-structure of the sentence:

(45) *continui*
$$
\begin{aligned}
&(f\ \textsc{pred}) = \text{`continue}\langle\textsc{subj}\rangle\text{'}\\
&(f\ \textsc{tense}) = \textsc{present}\\
&((f\ \textsc{subj pred}) = \text{`pro'})\\
&(f\ \textsc{subj number}) = \textsc{singular}\\
&(f\ \textsc{subj person}) \neq 3
\end{aligned}
$$

The negative equation for SUBJ PERSON in (45) correctly blocks the ungrammatical (44), while specifying no positive person information about the subject, which correctly reflects uncertainty of knowledge about the form (i.e., ambiguity). Another example of syncretism of agreement features is shown for English main verbs in Appendix A.

The f-description in (45) also demonstrates the standard LFG treatment of pro-drop: the verb optionally specifies that its SUBJ has the PRED value 'pro'. This allows the f-structure for a pro-drop sentence, such as (43), to satisfy Completeness, since the thematic SUBJ that the verb governs is present and has a PRED. The f-structure for (43), which satisfies the f-description (45), is:

(46)
$$
f\begin{bmatrix}
\textsc{pred} & \text{`continue}\langle\textsc{subj}\rangle\text{'}\\
\textsc{tense} & \textsc{present}\\
\textsc{subj} & \begin{bmatrix}\textsc{pred} & \text{`pro'}\\ \textsc{number} & \textsc{singular}\end{bmatrix}
\end{bmatrix}
$$

We noted above that multiple functional applications can be written as an f-structure label followed by a string of symbols, as in (f SUBJ NUMBER). Kaplan and Zaenen (1989) develop an f-structure-based theory of unbounded dependencies that relies on a simple extension to this, such that the string of symbols is drawn from a regular language. This means that optionality, negation, disjunction, complementation and Kleene star and plus are valid operations on the string of attributes in an f-structure constraint. The regular expression operators allow the statement of f-structure constraints that contain *functional uncertainty* and are thus resolvable in a (potentially unlimited) number of ways. This use of regular expressions is similar to the GPSG theory of unbounded dependencies, which is stated in terms of slash categories in phrase structure rules that support regular expressions (Gazdar 1981, Gazdar et al. 1985). One crucial difference, discussed below, is that the LFG functional uncertainty approach does not need to posit traces in phrase structure.

Let us consider an example. We noted in Table 17.1 that *wh*-phrases in interrogatives are assigned the discourse grammatical function FOCUS. Suppose that we want to allow the *wh*-phrase to correspond to the grammatical functions SUBJ or OBJ. We could then write the following equation:

(47) $(f\ \textsc{foc}) = (f\ \{\textsc{subj} \mid \textsc{obj}\})$

The right-hand side of the equation contains an uncertainty about which grammatical function the *wh*-phrase is identified with.

The equation in (47) does not yet capture the unbounded nature of *wh*-dependencies. Using the Kleene operators, we add a further, unbounded uncertainty over the grammatical functions in the f-structure that the dependency may licitly pass through. For example, the following equation states that the *wh*-dependency may pass through any number (including zero) of XCOMP or COMP grammatical functions and must be identified at the bottom of the dependency with a SUBJ or OBJ:

(48) $(f\ \textsc{foc}) = (f\ \{\textsc{xcomp} \mid \textsc{comp}\}^*\ \{\textsc{subj} \mid \textsc{obj}\})$

This captures the same effects as (47), but now allows for unboundedness, generating examples such as:

(49) Who saw this?

(50) What did John see?

(51) What did Mary say that John saw?

(52) What did Mary seem to say that John saw?

Island constraints and other constraints on extraction are captured through the path specification in the functional uncertainty equation. For example, the equation in (48) already captures the Sentential Subject Constraint, ruling out (53), because SUBJ is not on the extraction path: the dependency can terminate in a SUBJ, but cannot pass through one. Similarly, the equation captures the Left Branch Condition, ruling out (54), because the path cannot terminate in SPEC.

(53) * Who does [that John likes _] surprised Mary?

(54) * Whose did they steal [_ car]?

Equation (48) is just meant to be illustrative and does not capture the full range of grammatical possibilities nor rule out the full range of ungrammatical cases. What (48) shows, though, is that conditions on extraction are captured in LFG by appropriately limiting the extraction path, as expressed in a functional uncertainty equation. For a more complete specification of functional uncertainty paths, including pied-piping, see Dalrymple (2001: chapter 14). Some recent in-depth investigations of unbounded dependencies in LFG are Berman (2003), Asudeh (2004), and Mycock (2006). Berman (2003) and Asudeh (2004) consider the question of successive-cyclic effects in unbounded dependencies and consider an alternative to functional uncertainty based on functional control.

The LFG approach to unbounded dependencies that developed from Kaplan and Zaenen's functional uncertainty approach is notable in that it posits no traces or copies in the syntax—whether in c-structure or f-structure. See the appendix for an illustration. Bresnan (1995, 1998, 2001*b*) has argued from cross-linguistic data on weak crossover that traces are required in certain narrowly circumscribed circumstances, but see Dalrymple et al. (2001, 2007) for a traceless alternative and Berman (2003: chapter 5) for a critical appraisal of both sides of the debate.

The non-argument discourse functions FOCUS and TOPIC are subject to the following general principle (Zaenen 1980, Bresnan and Mchombo 1987):[7]

(55) **Extended Coherence Condition**
FOCUS and TOPIC must be linked to the semantic predicate argument structure of the sentence in which they occur through proper integration with the sentence's f-structure. Proper integration is either functional equality with or anaphoric binding of a grammatical function.

Functional equality is the integration mechanism that we have seen so far, which is appropriate for filler-gap dependencies. Anaphoric binding is appropriate for resumption, left-dislocation, hanging topics, and other phenomena in which the discourse function has a corresponding pronoun in the clause. See Asudeh (2004) for an in-depth treatment of resumption and discussion of related cases of satisfaction of the Extended Coherence Condition through anaphoric binding.

[7] Some formulations of the Extended Coherence Condition also apply to ADJUNCTS; see, for example, Bresnan (2001*b*: 63) and Falk (2001: 64).

The functional applications and functional uncertainties we have examined thus far have all been *outside-in*: in stating the constraint, some path is examined from an outer f-structure to an inner f-structure. The extension of the f-structure constraint language to allow functional uncertainty also enables *inside-out functional application* and *inside-out functional uncertainty* (first published in Halvorsen and Kaplan 1988), which permit constraints to be placed on paths from an inner f-structure to an outer f-structure.

Inside-out functional application is the formal foundation of the theory of *constructive case* developed by Nordlinger (1998) in her analysis of the Australian language Wambaya. In this theory, the case inflection directly determines the grammatical function of the nominal by stating which GF the nominal's f-structure must be the value of. We can demonstrate the generality of the idea by looking at an example from a typologically unrelated language, Malayalam (Mohanan 1982):

(56) Kuṭṭi aanaye aar̄aaḏiccu. Malayalam
child.NOM elephant.ACC worship.PAST
The child worshipped the elephant.

Mohanan (1982) notes that, in Malayalam, case-marking together with animacy determines the grammatical function of the nominal. For example, an animate nominative is a subject. This is captured through the following f-description that is part of the lexical information contributed by *kuṭṭi* ("child"), where f is the f-structure of the noun:

(57) *kuṭṭi* (f PRED) = 'child'
(f ANIMATE) = +
(f CASE) = NOMINATIVE
(SUBJ f)

The final constraint in (57) is an inside-out existential constraint which requires that there is an f-structure, call it g, such that the noun's f-structure is the value of g's SUBJ attribute. For formal definitions of outside-in and inside-out functional application and uncertainty, see Dalrymple (2001: 100–4, 143–6).

Inside-out functional uncertainty plays an important role in LFG's binding theory, as initially explored in Dalrymple (1993). Constraints on antecedents of anaphors are stated in f-descriptions according to the following general schema, where f is the f-structure of the anaphor:

(58) ((DomainPath f) AntecedentPath)

DomainPath is a path that states in which domain the antecedent of the anaphor must occur. It is stated in terms of an inside-out functional uncertainty relative to the f-structure of the anaphor. AntecedentPath then specifies where within this domain the antecedent may occur and which grammatical function the antecedent has. Dalrymple (1993) shows that this kind of equation, including constraints on properties of f-structures that DomainPath passes through, gives both a formally precise and typologically appropriate explanation of anaphoric binding possibilities.

As an example, let us consider the long-distance reflexive *aapaṇ* in Marathi, as discussed in Dalrymple (1993). This pronominal must be bound within the sentence, so it is an anaphor, but it cannot be bound locally (Dalrymple 1993: 14, 77):

(59) Tom mhanat hota ki Sue ni aaplyaalaa maarle. Marathi
Tom$_i$ said that Sue ERG self$_i$.ACC hit
Tom said that Sue hit him (Tom).

(60) *Jane ne aaplyaalaa bockaarle.
Jane ERG self.ACC scratched
Jane scratched herself.

The binding constraint on how *aapaṇ* is permitted to take an antecedent can be captured with the following inside-out functional uncertainty, where f is the f-structure of the reflexive:

(61) $((\text{GF}^{+}\ \text{GF}\ f)\ \text{GF})$

The specification of DomainPath as $(\text{GF}^{+}\ \text{GF}\ f)$ means that the antecedent is not in the f-structure of the reflexive, which is just $(\text{GF}\ f)$, but rather at least one further f-structure out (due to Kleene plus). This captures the fact that the reflexive cannot be bound locally. The AntecedentPath is simply GF, which allows the antecedent to bear any grammatical function, but this can be further restricted.

17.4 The C-structure to F-structure correspondence

We have now briefly looked at c-structure and f-structure and constraints on each kind of structure, but we have yet to explain how the two structures are related by structural correspondences. This section first explains how the mapping works, and then how LFG captures the empirical observation that radically different c-structures can correspond to the same f-structure: languages can express the same basic relation with strikingly different structural and morphological tools at their disposal.

17.4.1 How the c-structure to f-structure mapping works

The correspondence function ϕ maps c-structure nodes to f-structures. The mapping is deterministic (since it is a function) and many-to-one. The mapping is determined by language-specific instantiations of general mapping principles (Bresnan 2001*b*, Toivonen 2003) on annotated phrase structure rules. Lexical information is mapped from terminal nodes in c-structure, which contain all of the information lexically associated with the word. The annotations on c-structure nodes are functional constraints of the kind discussed in the previous section.

The mapping is stated in terms of two metavariables over f-structure labels, as defined in (62). These f-structure metavariables are defined in terms of a c-structure variable, $*$, which stands for "the current node", and the mother (i.e., immediate dominance) function on tree nodes, $\mathcal{M}$, where $\mathcal{M}(*)$ is "the node immediately dominating the current node". It is a common LFG convention to write $\widehat{*}$ instead of $\mathcal{M}(*)$.

(62) $\downarrow \equiv \phi(*)$
i.e., 'the f-structure of the current c-structure node' or 'my f-structure'
$\uparrow \equiv \phi(\widehat{*})$
i.e., 'the f-structure of the node that immediately dominates the current c-structure node' or 'my mother's f-structure'

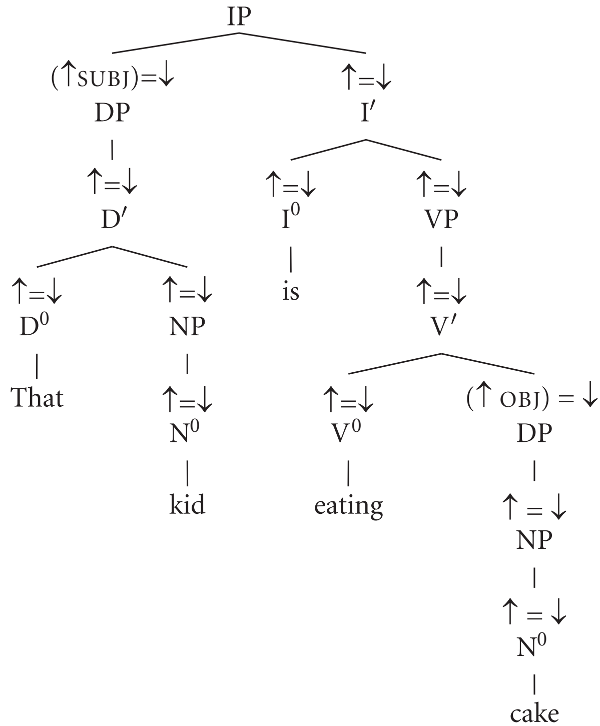
The up and down arrows are meant to symbolize their meaning graphically: since the annotations on non-terminals are typically written above the category label, the up arrow is pointing at the mother and the down arrow is pointing at the current node. This is essentially the original formalization of Kaplan and Bresnan (1982); see also Kaplan (1987, 1989). An alternative, strongly model-theoretic specification of the metavariables and LFG grammars more generally is provided by Blackburn and Gardent (1995).

The sample annotated phrase structure rule in (63) states that IP dominates a DP and an I′. The annotations specify that the information in I′ maps to the same f-structure as the information of its mother (the IP) and that the information contained in the DP maps into an f-structure that is the value of the SUBJECT grammatical function in the f-structure of the IP.

(63) $$\text{IP} \rightarrow \begin{array}{cc} \text{DP} & \text{I}' \\ (\uparrow \text{SUBJ}) = \downarrow & \uparrow = \downarrow \end{array}$$

The annotated version of the c-structure in (2) above, which presupposes a number of additional annotated phrase structure rules like (63), is given in (64). For presentational purposes, we henceforth suppress intermediate (bar-level) categories in non-branching sub-trees; this is common practice in the LFG literature.

(64)



The terminal nodes in c-structure are lexical entries, which specify the form of the word, its syntactic category, and a set of f-structure constraints (the lexical item's f-description). It is more strictly correct to write the f-description of the lexical item immediately below the word form in the c-structure, since the lexical item's f-description is actually part of the terminal node's information. However, for presentational reasons, we instead specify the lexical entries separately in (65):

(65) *that*, D^0 (↑ DEFINITE) = +
(↑ DEIXIS) = DISTAL
(↑ NUMBER) = SG
(↑ PERSON) = 3

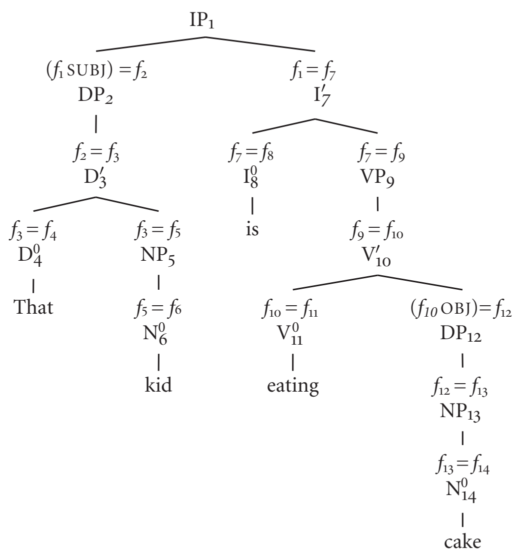
kid, N^0 (↑ PRED) = 'kid'
(↑ NUMBER) = SG
(↑ PERSON) = 3

is, I^0 (↑ SUBJ NUMBER) = SG
(↑ SUBJ PERSON) = 3
(↑ TENSE) = PRESENT
(↑ PARTICIPLE) $=_c$ PRESENT

$$
\begin{array}{ll}
\textit{eating}, \mathrm{V}^0 & (\uparrow \text{PRED}) = \text{`eat}\langle\text{SUBJ, OBJ}\rangle\text{'} \\
 & (\uparrow \text{ASPECT}) = \text{PROGRESSIVE} \\
 & (\uparrow \text{PARTICIPLE}) = \text{PRESENT} \\
\textit{cake}, \mathrm{N}^0 & (\uparrow \text{PRED}) = \text{`cake'} \\
 & (\uparrow \text{NUMBER}) = \text{SG} \\
 & (\uparrow \text{PERSON}) = 3
\end{array}
$$

The metavariables are instantiated as follows. Each c-structure node is assigned an arbitrary, unique index. The c-structure variable * for each node is instantiated as the node's index and the f-structure metavariable is instantiated accordingly. Up arrow metavariables in lexical f-descriptions are instantiated according to the label of the pre-terminal node that dominates the item in question. This should be intuitively clear if one bears in mind that the f-description is actually part of the terminal node. The instantiated version of (64) and its corresponding f-structure is shown in (66). Notice that we have adopted a typical convention of writing f_1 instead of $\phi(1)$ and so on.

(66)



$$
\begin{array}{l}
f_1 \\ f_7 \\ f_8 \\ f_9 \\ f_{10} \\ f_{11}
\end{array}
\left[
\begin{array}{lll}
\text{PRED} & & \text{`eat}\langle\text{SUBJ, OBJ}\rangle\text{'} \\
\text{SUBJ} & \begin{array}{l} f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{array} & \left[\begin{array}{ll} \text{PRED} & \text{`kid'} \\ \text{DEIXIS} & \text{DISTAL} \\ \text{DEFINITE} & + \\ \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{array}\right] \\
\text{OBJ} & \begin{array}{l} f_{12} \\ f_{13} \\ f_{14} \end{array} & \left[\begin{array}{ll} \text{PRED} & \text{`cake'} \\ \text{NUMBER} & \text{SG} \\ \text{PERSON} & 3 \end{array}\right] \\
\text{TENSE} & & \text{PRESENT} \\
\text{ASPECT} & & \text{PROGRESSIVE} \\
\text{PARTICIPLE} & & \text{PRESENT}
\end{array}
\right]
$$

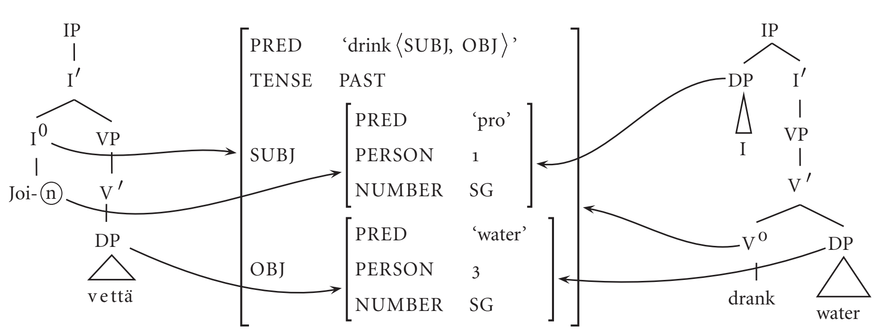
It should be noted that the features provided here reflect a specific analysis, and individual researchers may disagree on what the best analysis of a given phenomenon is. For example, we have treated the demonstrative *that* as just contributing features to the f-structure of the nominal head (*kid*). Others might propose that *that* projects to a SPEC f-structure and contains its own PRED.

17.4.2 Flexibility in mapping

The mappings between c-structure and f-structure and other structures are principled and unambiguous, based on the mechanisms presented in section 17.4.1.

However, there is cross-linguistic variation in exponence of linguistic information. For example, many languages rely more on morphology than hierarchical phrase structure in expressing syntactic information. This generalization is taken very seriously in LFG and is encapsulated in the slogan "morphology competes with syntax" (Bresnan 2001*b*: 6). Morphological information can be mapped directly into f-structure and there is thus no need to assume that all languages have the same, or similar, c-structure at some underlying level. In order to posit a highly articulated phrase structure for a given language, there must be evidence for such a structure. If a language expresses a grammatical function with a bound morpheme, the information is mapped directly from that morpheme onto the f-structure function: there is thus no need to posit an empty c-structure node for the grammatical function. Similarly, morphosyntactic information that is contributed by functional projections in other theories can be directly contributed morphologically in LFG.

Examples of cross-linguistic differences in c-structural expression abound. A pronominal subject may be expressed as an independent DP in some languages and a bound morpheme in others. Tense information is hosted by V^0 in some languages and I^0 in others, and in some languages it can be hosted by either I^0 or V^0. There is nothing about the mapping algorithm or the theory of c-structure that prohibits such c-structural differences between languages. Comparing two sentences with similar meanings in two different languages, the f-structures will look similar or identical and the c-structures may look radically different. Furthermore, f-structure information may be contributed simultaneously from different nodes in c-structure. In (67) we see an illustration of these points: the Finnish c-structure on the left side and the English c-structure on the right side map to the same f-structure:[8]

(67)


In sum, radically different c-structures may map to f-structures that are identical or near-identical.

[8] This is a slight oversimplification. F-structures expressing the same basic relations in two languages may contain certain differences. For example, languages can differ in the tense and aspect distinctions they make, whether they mark evidentiality, case marking, etc.

A language often has more than one way to express the same function. For example, Finnish has c-structurally independent subjects in addition to the morphologically bound pronominal subjects (compare examples (9) and (67)). Also, compare the two English examples in (68):

(68) a. Hanna poured out the milk.
b. Hanna poured the milk out.

The word *out* has the same basic function in (68a) and (68b). However, the phrase structural realization is different, as evidenced by the basic fact that the word order differs, but also by the observation that *out* in (68a) cannot have a complement or be modified, whereas *out* in (68b) can (for references and LFG analyses of the verbal particles in several Germanic languages, see Toivonen (2003)). The key point with respect to the examples in (68) is that their f-structural representation would be the same, while their c-structures differ.

The flexibility in mapping between c-structure and f-structure renders unnecessary highly abstract phrase structure representations that contain empty categories and functional projections hosting tense, aspect, case and other functional information. Instead, c-structural representations are faithful to the word order and constituency of the sentences they model. The theory of c-structure is very much a "what you hear is what you get" theory of surface syntax.

17.5 The correspondence architecture

The two structures, c-structure and f-structure, and correspondence function, ϕ, that we have examined so far constitute the original architecture of LFG, as laid out by Kaplan and Bresnan (1982). This architecture was subsequently generalized (Kaplan 1987, 1989, Halvorsen and Kaplan 1988) and the resulting architecture became known as the Parallel Projection Architecture or Correspondence Architecture.

The essential insight behind the Correspondence Architecture is that it is possible to resolve the apparent contradiction between, on the one hand, the empirically motivated proliferation of levels of representation and the resulting rich array of structures and constraints, and, on the other hand, formal elegance and theoretical parsimony (Kaplan 1987: 363). The resolution is accomplished as follows in the architecture. The notion of correspondence function is generalized from the ϕ function to include a number of other functions relating other structures. A rich set of structures and correspondences can be posited as constituting the linguistic form-meaning relation. However, since the correspondence functions are functions

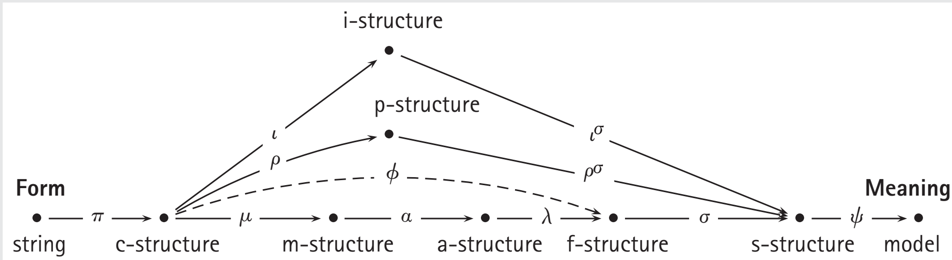


Figure 17.1. The Correspondence Architecture (Asudeh 2006)

in the mathematical sense, they can be composed into larger functions. Thus, despite the linguistic richness they offer, the correspondences are mathematically and computationally eliminable (Kaplan 1987: 363).

Kaplan (1987, 1989) suggests a programmatic version of the architecture, but the first theoretically well-developed version of the architecture added semantic structure (abbreviated alternatively as s-structure or sem-structure; Halvorsen and Kaplan 1988, Dalrymple 1993). Semantic structure then formed part of the basis for Glue Semantics, a theory of the syntax-semantics interface and semantic composition (Dalrymple et al. 1993, Dalrymple 1999, 2001, Asudeh 2004, Lev 2007, Kokkonidis 2008). Glue Semantics has become the predominant semantic theory for LFG, but is actually an independent theory that could in principle be integrated with other syntactic theories; for example, Asudeh and Crouch (2002) define Glue Semantics for Head-Driven Phrase Structure Grammar and Frank and van Genabith (2001) for Lexicalized Tree-Adjoining Grammar. Asudeh (2006) considers questions of semantic composition in light of the Correspondence Architecture and provides a version of the architecture that incorporates numerous proposals in the LFG literature subsequent to the addition of semantic structure. Asudeh's presentation of the Correspondence Architecture is shown in Figure 17.1.

Let us examine this version of the architecture briefly. There is an explicit correspondence, π, between the string and the c-structure, as proposed by Kaplan (1987, 1989). An alternative theory of the string to c-structure mapping is pursued by Wescoat (2002, 2005, 2007) in a theory of *lexical sharing* that defines a way for words to project to more than one terminal node (interestingly, without changing the formal status of c-structures as trees). Information structure (i-structure; Butt and King 2000*b*) encodes notions like discourse topic and focus and old and new information. Phonological structure (p-structure; Butt and King 1998*a*, O'Connor 2006) models phrasal phonology and prosody (a more accurate name might in fact be prosodic structure). Mycock (2006) uses p-structure in her analysis of *wh*-in-situ phenomena, which uses the correspondence architecture to account for these in terms of p- and f-structure rather than positing covert movement or empty c-structure positions. Information structure and phonological structure have both been proposed as projections from c-structure. Argument structure (a-structure; Butt et al. 1997) has been proposed for modeling semantic role information. Morphological structure (m-structure; Butt et al. 1996, 1999, Frank and Zaenen 2002, Sadler and Spencer 2004) has been proposed as an interface between syntax and morphology to capture in a more principled manner information that might otherwise be included in f-structure (e.g., tense-aspect information from auxiliaries). There has been some debate over the proper location for m-structure in the architecture. Butt et al. (1996, 1999) treat it as a projection from c-structure. Frank and Zaenen (2002) argue that, although this is adequate for the phenomena for which Butt et al. (1996, 1999) use morphological structure (auxiliaries), there are reasons to prefer morphological structure as a projection from f-structure. We assume, following Asudeh (2006), that morphological information should feed both argument structure and functional structure and therefore place m-structure between c-structure and a-structure. The resulting architecture demonstrates Kaplan's point about function composition. The original ϕ function of Kaplan and Bresnan (1982) is the composition of μ, α and λ : $\phi = \lambda \circ \alpha \circ \mu$. Lastly, we note that the mapping ψ from semantic structure to meaning is assumed to be characterized by proofs in Glue Semantics; see Asudeh (2006) for more details.

17.6 Some recent developments

Optimality-Theoretic LFG (OT-LFG) is a relatively recent outgrowth of the theory that uses LFG as the GEN component in an Optimality Theory (OT; Prince and Smolensky 1993, 2004) syntax. Parts of the constraints in the EVAL component in OT-LFG are also stated using formal notions from LFG. This extension of LFG was launched by Bresnan (1997, 2000, 2001*a*). It has been pursued in numerous publications in the proceedings of the annual LFG conference. Some other major works on OT-LFG are Morimoto (2000), Lee (2001), Sells (2001*a*,*b*), Kuhn (2003), and Clark (2004). An interesting recent development has seen OT-LFG applied to explaining dialect variation (Bresnan et al. 2007). Lastly, Optimality Theory has also influenced computational work on LFG, where the OT-inspired notion of *optimality marks* (Frank et al. 1998) is used for robustness of parsing and control of generation (Butt et al. 1999: 199–204). However, this latter application of OT stops short of OT-LFG's tight integration of the two theories; rather, a simple OT-inspired preference mechanism is overlaid on an LFG grammar to guide the grammar's parsing and generation.

Computational work on LFG continues to be a vital research area. There are several noteworthy research programs; here we identify just three. The Parallel Grammar project (ParGram; http://www2.parc.com/isl/groups/nltt/pargram/) is a collaborative international effort that seeks to develop implemented wide coverage LFG grammars based on a common inventory of f-structure features, with the goal of ensuring substantial commonality of f-structures (Butt et al. 1999, 2002). This collaborative activity not only has the consequence of testing and developing typological aspects of LFG, it also provides important insights and resources for machine translation. A recent offshoot of ParGram is the Parallel Semantics project (ParSem), which seeks to develop semantic structures for the grammars in the ParGram project. ParSem is strongly influenced by the second computational trend: inference of semantic representations from f-structures. This approach to semantics is often called Transfer Semantics, because the aim is to transfer relevant predicate-argument relations encoded in informationally rich "packed f-structures" to (packed) semantic representations in a computationally efficient manner (Crouch 2005, 2006, Crouch and King 2006). Transfer Semantics is an important component in industrial applications, such as the Powerset search engine. A third trend in computational work is research on automatic induction of LFG grammars (Cahill et al. 2005, O'Donovan et al. 2005, Cahill et al. 2008).

17.7 Concluding remarks

LFG differs from other syntactic theories in its adoption of formally and conceptually distinct syntactic structures (c-structure and f-structure). Although Relational Grammar has a structure that is similar to f-structure in that it models grammatical functions, it does not articulate a theory of constituent structure. Head-Driven Phrase Structure Grammar represents constituency and grammatical functions—indeed, all grammatical information—in a single formal structure. Principles and Parameters Theory does not acknowledge grammatical functions as such at all, attempting to derive them from phrase structure, which is the representation used to model all syntactic information.

In addition to grammatical modularity, another underlying principle of LFG theory is that grammatical information grows *monotonically* (Bresnan 2001*b*: chapter 5), i.e., in an information-preserving manner. For example, as an f-description grows in size through the addition of new defining equations, the minimal f-structure that models the description also grows in size, becoming increasingly specific. Addition of constraining equations and other constraints similarly does not remove information but rather constrains the existing minimal model. Growth of an f-description never results in information loss. This has a number of further consequences. One general consequence is that there can be no destructive operations in syntax. For example, relation-changing operations, such as passive, cannot be syntactic, because that would require destructive remapping of grammatical functions. Another general consequence is that grammatical information of parts of linguistic expressions are preserved in the grammatical information of the whole. This in turn means that the parts can form informative fragments (Bresnan 2001*b*: 79–81). Fragments are an important part of LFG's robustness for computational parsing, since parts of ungrammatical sentences are often grammatical, and these grammatical parts can be returned in a set of wellformed fragments (Crouch et al. 2008). Cognitive aspects of fragments have also been explored, in a psycholinguistic model of human parsing and production (Asudeh 2004: chapter 8).

LFG is unique in its popularity both among computational linguists, who investigate and capitalize on formal and algorithmic properties of LFG grammars, and among descriptive and documentary linguists, who use the theory as a tool to understand and document understudied languages. We have already mentioned some of the research in computational linguistics and grammar engineering that relies on and develops LFG grammars and theory. LFG's usefulness for language description is summarized aptly by Kroeger (2007):

> LFG has a number of features that make it an attractive and useful framework for grammatical description, and for translation. These include the modular design of the system, the literal representation of word order and constituency in c-structure, a typologically realistic approach to universals (avoiding dogmatic assertions which make the descriptive task more difficult), and a tradition of taking grammatical details seriously. (Kroeger 2007: 1)

Last, but not least, the third group of researchers who have adopted LFG are traditional theoretical linguists. The characteristics that Kroeger lists above are also useful for theoretical analysis and have resulted in substantial insights into natural language. Also, many theoretical linguists find it useful that there are computational tools available to implement and test new theoretical claims. This is further facilitated by the fact that the major computational implementation, the XLE grammar development platform (Crouch et al. 2008), reflects LFG theory directly. In other words, the implementation and the theory are congruent, rather than the XLE implementing some ad hoc version of the theory.

The correspondence architecture of LFG has also proven useful for purposes that the main architects perhaps had not anticipated. For example, it offers an excellent framework for analyzing historical change (Vincent 2001). The framework allows us to pose and answer questions such as: What is the nature of the change: Is the change morphological? C-structural? F-structural? Does the change concern a specific type of linguistic information, or does the change concern the mapping between different types of information? A further advantage of LFG is its explicit and detailed representation of lexical information as lexical features. A small change in lexical information can have major syntactic consequences. Thus, both synchronic and diachronic variation can be readily represented as lexical variation. LFG has been used to model historical change by Allen (1995), and others (see, for example, the collection of papers in Butt and King 2001*b*).

Further Resources

Dalrymple (2001) is a standard reference work on LFG that reviews and develops the formal theory in considerable detail against a wide-ranging empirical backdrop. Bresnan (2001*b*) is an advanced textbook on LFG that also introduces certain theoretical innovations; the second edition is currently in preparation (Bresnan et al., in prep.). Two introductory textbooks are Falk (2001) and Kroeger (2004). Butt et al. (1999) is an introduction to grammar engineering with LFG grammars in XLE, although there have been many subsequent developments since its publication. The authoritative source for the Xerox Linguistic Environment is the included documentation (Crouch et al. 2008). XLE is not currently open source or freely available, but a free educational license may be obtained from the NLTT group at PARC. Bresnan (1982*c*), Dalrymple et al. (1995) and Butt and King (2006) are collections of many of the seminal early papers on LFG. Numerous monographs and edited volumes on LFG are published by CSLI Publications, who also publish online the proceedings of the annual LFG conference (http://cslipublications.stanford.edu/site/ONLN.shtml); the proceedings are freely available. Lastly, there is an LFG web page that serves as a general portal (http://www.essex.ac.uk/linguistics/external/LFG/).

Appendix A example: unbounded dependency, adjuncts, raising, control

(69) What did the strange, green entity seem to try to quickly hide?

(70) **Lexicon**

what, D^0	($\uparrow$ PRED) = ‘pro’
	($\uparrow$ PRONTYPE) = WH
	((FOCUS $\uparrow$) MOOD) $=_c$ INTERROGATIVE
did, I^0	($\uparrow$ TENSE) = PAST
	($\uparrow$ MOOD) = DECLARATIVE
	($\uparrow$ VFORM) $=_c$ BASE
	$\vee$
C^0	($\uparrow$ TENSE) = PAST
	($\uparrow$ MOOD) = INTERROGATIVE
	($\uparrow$ VFORM) $=_c$ BASE
the, D^0	($\uparrow$ DEFINITE) = +
strange, A^0	($\uparrow$ PRED) = ‘strange’

green, A^0	$(\uparrow \text{PRED}) = \text{‘green’}$
entity, N^0	$(\uparrow \text{PRED}) = \text{‘entity’}$
	$(\uparrow \text{NUMBER}) = \text{SG}$
	$(\uparrow \text{PERSON}) = 3$
seem, V^0	$(\uparrow \text{PRED}) = \text{‘seem}\langle\text{XCOMP}\rangle\text{SUBJ’}$
	$(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$
	$\{\neg[(\uparrow \text{SUBJ NUMBER}) = \text{SG}$
	$(\uparrow \text{SUBJ PERS}) = 3]$
	$(\uparrow \text{TENSE}) = \text{PRESENT}\,\vert$
	$(\uparrow \text{VFORM}) = \text{BASE}\}$
to, I^0	$\neg(\uparrow \text{TENSE})$
try, V^0	$(\uparrow \text{PRED}) = \text{‘try}\langle\text{SUBJ, XCOMP}\rangle\text{’}$
	$(\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})$
	$\{\neg[(\uparrow \text{SUBJ NUMBER}) = \text{SG}$
	$(\uparrow \text{SUBJ PERS}) = 3]$
	$(\uparrow \text{TENSE}) = \text{PRESENT}\,\vert$
	$(\uparrow \text{VFORM}) = \text{BASE}\}$
quickly, Adv^0	$(\uparrow \text{PRED}) = \text{‘quickly’}$
hide, V^0	$(\uparrow \text{PRED}) = \text{‘hide}\langle\text{SUBJ, OBJ}\rangle\text{’}$
	$\{\neg[(\uparrow \text{SUBJ NUMBER}) = \text{SG}$
	$(\uparrow \text{SUBJ PERS}) = 3]$
	$(\uparrow \text{TENSE}) = \text{PRESENT}\,\vert$
	$(\uparrow \text{VFORM}) = \text{BASE}\}$

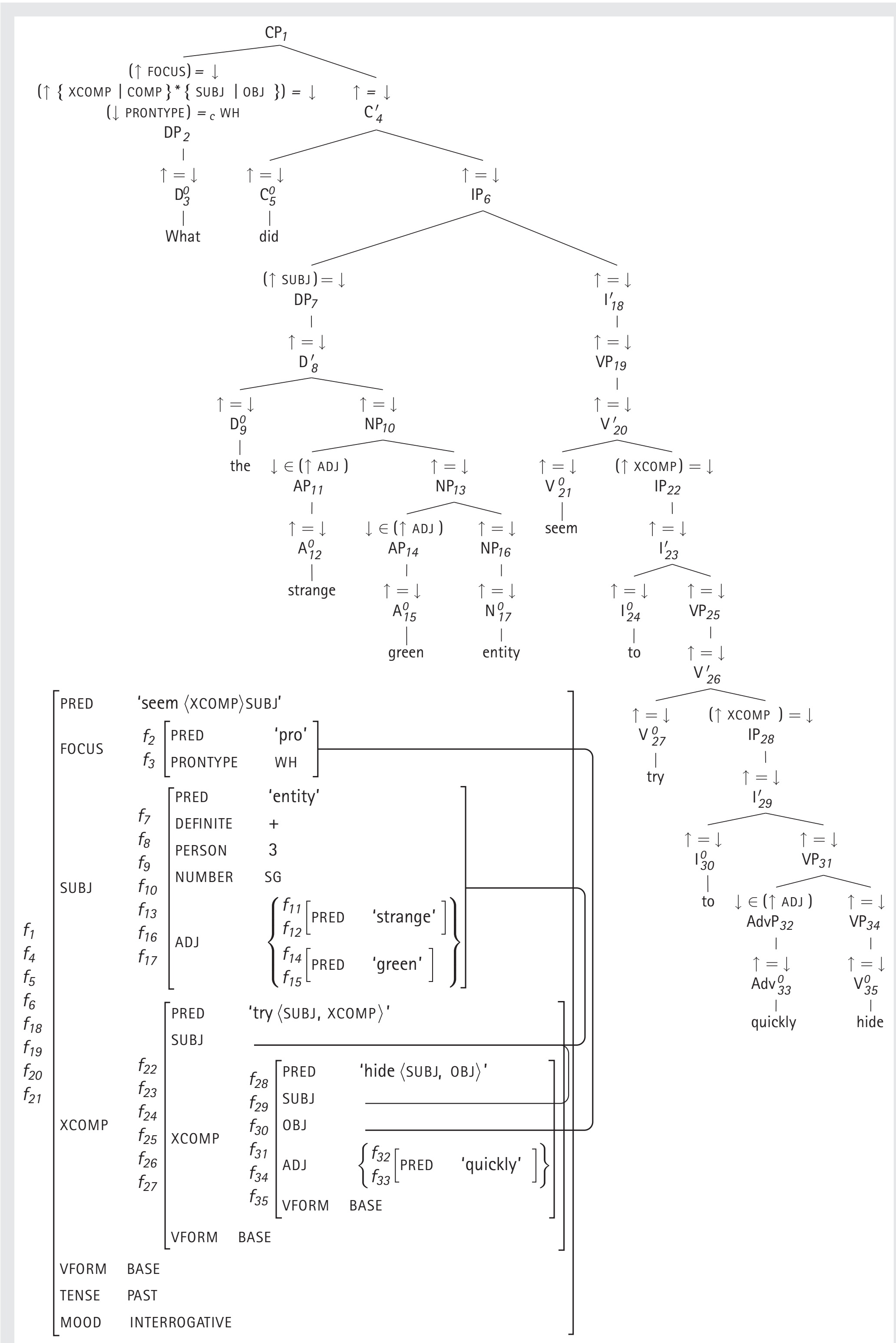


Figure 17.2. C-structure and f-structure for (69)

CHAPTER 18

THE NATURAL SEMANTIC METALANGUAGE APPROACH

CLIFF GODDARD

18.1 Goals, assumptions, and priorities

The basic conviction behind the NSM approach—bolstered by scores of empirical studies—is that meaning is the key to insightful and explanatory descriptions of most linguistic phenomena, phonetics and phonology excepted. Meaning is also the bridge between language and cognition, and between language and culture. Compartmentalizing language (or linguistic analysis) into syntax, morphology, semantics, and pragmatics therefore makes little sense. In linguistics, meaning is everybody's business.

The Natural Semantic Metalanguage (NSM) is a decompositional system of meaning representation based on empirically established universal semantic primes, i.e., simple indefinable meanings which appear to be present as word-meanings in all languages (Wierzbicka 1996*a*; Goddard 1998; Goddard and Wierzbicka 2002*b*; Peeters 2006; Goddard 2008). Originating with Wierzbicka

(1972), the system has been developed and refined over some 35 years. There is a large body of descriptive-analytical work in the framework, not only about English but Russian, Polish, French, Spanish, Malay, Japanese, Chinese, Korean, Ewe, East Cree, and other languages.[1] In addition to Anna Wierzbicka, Cliff Goddard, Jean Harkins, Bert Peeters, Felix Ameka, and other "old hands", there is a raft of new generation NSM researchers, such as Catherine Travis, Rie Hasada, Marie-Odile Junker, Uwe Durst, Kyung-Joo Yoon, Zhengdao Ye, Jock Wong, Anna Gladkova, Adrian Tien, and Helen Bromhead.

The NSM approach grew out of lexical semantics, and it still accords much greater importance to lexical semantics than many rival approaches, but the approach has long since extended into grammatical and illocutionary semantics, and (with the theory of cultural scripts) into cultural pragmatics. Adopting a uniform method of meaning description across these domains allows for the integration of areas of linguistic structure that often appear disparate and disconnected in other models.

The NSM metalanguage can be thought of as a highly disciplined and standardized subset of natural language: a small subset of word-meanings (63 in number, see Table 18.1 and also Appendix A), together with a subset of their associated grammatical properties. The NSM metalanguage itself represents a very substantial set of claimed findings about language universals: the existence of a specifiable set of fundamental lexico-semantic primes shared by all languages, with their shared combinatorial (syntactic) properties constituting a universal grammar. In tandem with this claim about linguistic universals there is a corresponding claim about universals of cognition, because the mini-language of semantic primes embodies the fundamentals of linguistic cognition, i.e., cognition as it can be carried out with and expressed through language (Whorf 1956). The NSM metalanguage is not just, however, an object of study in its own right. It is an invaluable descriptive tool for the analysis and contrastive study of meaning-related phenomena in all languages: a *tertium comparationis* for cross-linguistic study and language typology.

The attractions of an approach to meaning representation based on simple word-meanings in ordinary language can be itemized as follows. First, any system of representation has to be interpreted in terms of some previously known system, and since the only such system shared by all language users is natural language itself, it makes sense to keep the system of semantic representation as close as possible to natural language. Second, clear and accessible semantic representations enhance the predictiveness and testability of hypotheses. Most other systems of semantic analysis are hampered by the obscurity and artificiality of the terms of description. Third,

[1] A bibliography of NSM publications, along with a number of downloadable papers, is available at the NSM Homepage [www.une.edu.au/bcss/linguistics/nsm].

the system is intended to represent the cognitive reality of ordinary language users, so it would be problematical to employ symbols whose meanings are completely opaque to language users themselves.

The formal mode of meaning representation in the NSM approach is the semantic explication. This is a reductive paraphrase—an attempt to say in other words (in the metalanguage of semantic primes) what a speaker says when he or she utters the expression being explicated. As far as I am aware, NSM is the only approach to employ paraphrase in a strict sense. Many systems seek to describe meaning in decompositional terms, but decompositional or not, there is an enormous difference between paraphrase and description. For one thing, paraphrase attempts to capture an insider perspective (with its sometimes naïve first-person quality, rather than the sophisticated outsider perspective of an expert linguist, logician, etc.). Equally, paraphrase requires formulation in terms which are accessible and intelligible to the speakers concerned. The ready intelligibility of NSM explications to native speakers and the close relationship between the metalanguage and the language being described makes it easy to generate and test hypotheses: by direct or indirect substitution into natural contexts of use, and by direct accessibility to native speaker intuition.

A distinctive aspect of the NSM approach is the close attention it pays to the metaterminology of grammatical description, and in particular the need to achieve greater clarity and greater consensus about the meanings and operational criteria for grammatical terms such as "agent", "dative", "causative", "relative clause", "adverbial clause", and so on. The NSM approach seeks to identify for each such term a semantic prototype which can be used as a standard for the cross-linguistic identification of constructions of a given kind. In this way, the practice of linguistic typology can be "anchored" in semantic terms. Within a single language, NSM research indicates that any given grammatical construction is likely to be polysemous, i.e., to constitute a family of interrelated lexico-grammatical constructions with a prototype-plus-extensions structure.

Whether or not one is convinced of the universality of the NSM primes, many linguists would agree that there is heuristic value in plain language paraphrase in terms of a small standardized set of simple words. The long shelf life of many NSM studies would seem to confirm this. Furthermore, analyses framed in plain language paraphrase are available for later re-formulation in more technical terms, if required. Notwithstanding its simple mode of representation, the NSM approach has developed a fairly sophisticated suite of theoretical constructs, including the following, which will be explained and illustrated in the main body of this chapter: semantic primes, allolexy, syntactic frames and valency options of primes, semantic templates, semantic molecules, semantic prototypes for grammatical constructs, grammatical polysemy, ethnosyntax, and cultural scripts.

Table 18.1. Semantic primes, grouped into related categories

I, YOU, SOMEONE, SOMETHING/THING, PEOPLE, BODY	substantives
KIND, PART	relational substantives
THIS, THE SAME, OTHER/ELSE	determiners
ONE, TWO, SOME, ALL, MUCH/MANY	quantifiers
GOOD, BAD	evaluators
BIG, SMALL	descriptors
KNOW, THINK, WANT, FEEL, SEE, HEAR	mental predicates
SAY, WORDS, TRUE	speech
DO, HAPPEN, MOVE, TOUCH	actions, events, movement, contact
BE (SOMEWHERE), THERE IS, HAVE, BE (SOMEONE/SOMETHING)	location, existence, possession, specification
LIVE, DIE	life and death
WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT	time
WHERE/PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE	space
NOT, MAYBE, CAN, BECAUSE, IF	logical concepts
VERY, MORE	intensifier, augmentor
LIKE/WAY	similarity

Notes: (i) Primes exist as the meanings of lexical units (not at the level of lexemes) (ii) Exponents of primes may be words, bound morphemes, or phrasemes (iii) They can be formally complex (iv) They can have combinatorial variants (allolexes) (v) Each prime has well-specified syntactic (combinatorial) properties.

18.2 Universal semantic core

The inventory of semantic primes is listed (using English exponents) in Table 18.1. They are simple and intuitively intelligible meanings grounded in ordinary linguistic experience. The exponents of primes can be formally complex. The English words SOMEONE and SOMETHING, for example, consist of two morphological elements, and A LONG TIME and FOR SOME TIME are phrasemes. The NSM claim is that these expressions each represent unitary meanings. Not surprisingly, in many languages their equivalents are morphologically simple. Exponents of primes can also have multiple realizations (allolexes) in a single language. The "double-barreled" items in Table 18.1, such as SOMETHING/THING and OTHER/ELSE, indicate meanings which, in English, are expressed by means of different allolexes in different grammatical contexts. *Something* and *thing*, for example, express the same meaning, except that *something* is not normally used in combination with a specifier. Compare (a) *Something happened*, (b) *The same thing happened again*, (c) *I don't know when this thing happened*. Patterns of allolexy can vary from language to language.

Semantic primes exist, not at the level of whole lexemes, but as the meanings of lexical units. Language-specific polysemy can therefore obscure the identification of individual primes. A great deal of empirical research exists in the NSM literature on how semantic primes manifest themselves across languages. In particular, "whole metalanguage" studies have been carried out for English, Amharic, Polish, Russian, Malay, Lao, Mandarin, Mbula, Spanish, Korean, and East Cree, and more selective studies on French, Italian, Japanese, Bunuba, Ewe, Yankunytjatjara, Hawaiian Creole English, among others. On the basis of these studies, semantic primes appear to be lexical universals in the sense of having an exact translation in every human language (though the term "lexical" is used here in a broad sense, since it includes phrasemes and bound morphemes, as well as words proper).

It is not possible here to account in detail for how the primes were identified in the first place, but an example may be helpful. Consider the word *say*, in sentences like *Mary said something to me*. How could one paraphrase the meaning of *say* in this context, using simpler words? An expression like *verbally express* will not do, because terms like *verbally* and *express* are more complex and difficult to understand than *say* is in the first place. The only plausible line of explication would be something like 'Mary did something, because she wanted me to know something'; but this fails because there are many actions a person could undertake because of wanting someone to know something, aside from saying. On the other hand, if one takes a word like *ask*, as in *Mary asked me something*, it seems readily paraphrasable in simpler terms, including SAY, WANT, and KNOW: 'Mary said something to me, because she wanted to know something; she wanted me to say something because of this.' On account of its resistance to paraphrase, SAY is a good candidate for the status of semantic prime. Furthermore, SAY is clearly required for the explication of many other lexical items involving speaking and communication, especially speech-act verbs, as well as many discourse particles. Upon checking in a range of languages, one finds that all languages appear to have a word with the same meaning as English *say*. For example: Malay *kata*, Yankunytjatjara *wangkanyi*, Japanese *iu*.

As mentioned, polysemy is frequently a complication when trying to identify primes and match them up across languages. Often the range of use of exponents of the same prime do not coincide because, aside from the identical shared meaning, the words in question also have additional meanings which differ from language to language. After some 15 years of research, NSM researchers have accumulated a lot of data about common patterns of polysemy. Some widely attested patterns are summarized in Table 18.2. In NSM studies language-specific evidence is always adduced to support claims for semantic primes which depend on a polysemy analysis.

A complete outline of the natural semantic metalanguage of course calls for a specification of its grammar, as well as its lexicon, but we will defer this till section 18.4, and proceed straight to lexical semantics.

Table 18.2. Some common polysemies involving exponents of semantic primes

SAY	'speak', 'make sounds'	Thai, Mandarin, Yankunytjatjara, Kalam
THINK	'worry', 'long for', 'intend'	Mandarin, Swedish
WANT	'like', 'love'	Spanish, Ewe, Bunuba
HAPPEN	'arrive', 'appear'	French, Ewe, Mangaaba-Mbula
DO	'make'	Spanish, Malay, Arrernte, Samoan, Kalam, Amharic
BEFORE	'first', 'go ahead', 'front'	Lao, Samoan, Kayardild, Ewe, Mangaaba-Mbula
FEEL	'taste', 'smell', 'hold an opinion'	Malay, Acehnese, Ewe, French, Mandarin
WORDS	'what is said, message', 'speech, language'	Yankunytjatjara, Korean, Mangaaba-Mbula, Malay

Source: studies in Goddard and Wierzbicka 1994; 2002*b*; Goddard 2008.

18.3 Lexical semantics

There is a large body of descriptive empirical work in the NSM framework on many aspects of lexical semantics, with literally hundreds of published explications. Some lexicon areas that have been explored in great depth are emotions and other mental states, speech acts, causatives, cultural values, natural kind words, concrete objects, physical activity verbs, and discourse particles.

Doing NSM analysis is a demanding process and there is no mechanical procedure for it. Published explications have often been through a dozen or more iterations over several months. The validity of NSM explications can be tested on the basis of two main conditions. The first is substitutability in a broad sense: explications have to make intuitive sense to native speakers when substituted into their contexts of use, and to generate the appropriate entailments and implications. The second condition is well-formedness: they have to be framed entirely in semantic primes or molecules, and to conform to the syntax of the natural semantic metalanguage. In addition, explications have to conform to a coherence condition, i.e., they have to make sense as a whole, with appropriate chains of anaphora, co-reference, causal links, etc. In relation to the substitutability condition, NSM semantics makes extensive use of linguistic corpora, and (more recently) of internet searches using the Google search engine (though these have to be undertaken with due caution).

Over the 35 years since Wierzbicka's (1972) *Semantic Primitives*, the NSM program has developed new models of semantic explication capable of representing remarkable semantic detail and complexity. The "look and feel" of NSM explications can be illustrated with a series of thumbnail examples from English: first, with words that can be explicated directly in terms of semantic primes, and then with more complex examples that call for the use of intermediate-level "semantic molecules".

18.3.1 Explicating directly into semantic primes

Verbs *kill* and *break*. The causative verbs *kill* and *break* are frequently analyzed in the general linguistic literature as CAUSE TO DIE (or, CAUSE TO BECOME NOT ALIVE) and CAUSE TO BECOME BROKEN, respectively. NSM explications are given below. Aside from the fact the NSM system recognizes BECAUSE (rather than CAUSE) as its basic exponent in the causal domain, it can be seen that the explications give a more articulated and nuanced account of the event structure. In both cases, the explications depict an action by the agent X with an immediate effect on the patient Y, and, consequently the cessation of a prior state which otherwise would have continued. In the case of *kill*, an intermediate event is also involved, namely, something happening to person Y's body.

[A] *Someone X killed someone Y*:

someone X did something to someone else Y
because of this, something happened to Y at the same time
because of this, something happened to Y's body
because of this, after this Y was not living anymore

Break is both more complex than *kill*, and more polysemous. The explication below applies only to one sense of the word, as found in examples like *to break* a stick, an egg, a lightbulb, a vase, or a model plane. There is an aspectual component, namely, that the immediate effect on thing Y 'happened in one moment', and a final "subjective" component indicating that the result (i.e., 'Y was not one thing anymore') is seen as irrevocable or irreversible. It is an interesting fact, and one consistent with the somewhat schematic nature of this explication, that many languages lack any comparably broad term which would subsume many different manners of "breaking" (Majid and Bowerman 2007).

[B] *Someone X broke something Y*:

someone X did something to something Y
because of this, something happened to Y at the same time
it happened in one moment
because of this, after this Y was not one thing anymore
people can think about it like this: "it can't be one thing anymore"

Adjectives *sad* and *unhappy*. According to NSM research, the meanings of emotion terms involve feelings linked with a characteristic or prototypical cognitive scenario involving thoughts and wants (Wierzbicka 1999; Harkins and Wierzbicka 2001). The scenario serves as a kind of "reference situation" by which the nature of the associated feeling can be identified. For example, *joy* is linked prototypically with the thought 'something very good is happening now', *remorse* is linked with the thought 'I did something bad'. Consider the explication for English *to feel sad*.

[C] *X felt sad*:

someone X felt something bad
someone can feel something like this when this someone thinks like this:
"I know that something bad happened
I don't want things like this to happen
I can't think like this: I will do something because of it now
I know that I can't do anything"

The prototypical cognitive scenario involves an awareness that 'something bad happened' (not necessarily to me) and an acceptance of the fact that one can't do anything about it. This is compatible with the wide range of use of *sad*; for example, that I may feel *sad* when I hear that my friend's dog died, or when I think about some unpleasant bickering in my workplace.

This format of explication enables subtle meaning differences to be modeled across languages and within a single language. Consider some of the ways in which being *unhappy* differs from being *sad*: (i) Being *unhappy* requires the experiencer to have certain real thoughts (while one can say *I feel sad, I don't know why*, it would be a little odd to say *I feel unhappy, I don't know why*). (ii) *Unhappy* conveys a stronger negative evaluation, as implied by the fact that it is less readily combinable with minimizing qualifiers like *a little* or *slightly*. (iii) *Unhappy* has a more personal character: one can be saddened by bad things that have happened to other people, but if one is *unhappy*, it is because of bad things that have happened to one personally. (iv) *Unhappy* does not suggest a resigned state of mind but rather focuses on some thwarted desires. The attitude is not exactly active, because one doesn't necessarily want anything to happen, but it is not passive either. (v) *Unhappy* suggests a state extended in time. All these differences are modeled in the differences between the two explications.

[D] *X felt unhappy*:

someone X felt something bad
someone can feel something like this when this someone thinks like this for some time:
"some very bad things happened to me
I wanted things like this not to happen to me
I can't not think about it"
this someone felt something like this
because this someone thought like this

18.3.2 Semantic molecules

According to NSM research, some kinds of concept (emotions, values, speech acts, and interpersonal relations) are semantically much simpler than others (artifacts, animals and plants, and many human activities), because the former can be explicated directly in terms of semantic primes, while the latter can only be explicated in stages using intermediate-level semantic molecules. For example, the concept of 'animal' is necessary in the explications of *cat*, *mouse*, *dog*, *horse*, etc.; body-part concepts are required in verbs like *eat*, *punch*, and *run*; and almost all concrete vocabulary items require concepts such as 'long', 'round', 'flat', 'hard', among others.

A semantic molecule is a packet of semantic components which exists as the meaning of a lexical unit. Semantic molecules have a special cognitive significance in that they allow a conceptual chunking which makes it possible to manage concepts of great semantic complexity. It must be said immediately that there are many recurrent components that are not semantic molecules, because they are not encapsulated as the meanings of lexical items. For example, top-level categorical components for nouns such as 'one part of someone's body' (for body-part terms), 'living things of one kind' (for natural kind terms), and high-level verb components related to semantic roles, such as 'something happened in a place', 'someone did something', and 'something happened to something else because of it'. Such recurrent components can be extremely significant for the interface between lexical and grammatical semantics, and for the creation of lexical classes, but they are simple enough to be spelled out in relatively short strings composed purely of primes.

Now consider these examples of body-part words (Wierzbicka 2007*a*). The notation [M] indicates a semantic molecule. The claim is that *head* (in the sense of a human person's *head*) requires the shape descriptor 'round [M]', and that words like *legs*, *arms*, and *tail* require 'long [M]'.

[E] *head (someone's head)*:

one part of someone's body
this part is above all the other parts of the body
this part is like something round [M]
when someone thinks about something, something happens in this part of this someone's body

[F] *legs (someone's legs)*:

two parts of someone's body
these two parts are below all the other parts of the body
these two parts are long [M]
these two parts of someone's body can move as this someone wants
because people's bodies have these parts, people can move in many places as they want

It would be incorrect to assume that shape descriptors are more basic than all body-part terms, however, because one human body-part, namely *hands*, is necessary in the explication of shape descriptors themselves. This is because shape descriptors designate properties which are both visual and "tangible", and to spell out the nature of the latter concept requires the semantic prime TOUCH (contact) and the semantic molecule 'hands [M]'. For example:

[G] *something long (e.g., a tail, a stick, a cucumber)*:

when someone sees this thing, this someone can think about it like this:
"two parts of this thing are not like any other parts,
because one of these two parts is very far from the other"
if someone's hands [M] touch this thing everywhere on all sides,
this someone can think about it in the same way

From an experiential point of view the importance of the semantic molecule 'hands [M]' is perhaps not surprising. The experience of "handling" things, of touching them with one's hands and moving the hands in an exploratory way plays a crucial role in making sense of the physical world, and in our construal of the physical world. It turns out that, unlike many other body-part words, 'hands' can be explicated directly in terms of semantic primes, though space prevents us demonstrating this here (Wierzbicka 2007*b*: 47).

How many productive semantic molecules are there? At the current early stage of research, the answer is not very clear. For English, perhaps 150–250. It is known that productive molecules in English are drawn from at least the following categories (examples given are non-exhaustive): (a) parts of the body: 'hands', 'mouth', 'legs'; (b) physical descriptors: 'long', 'round', 'flat', 'hard', 'sharp', 'straight'; (c) physical activities: 'eat', 'drink', 'sit'; (d) physical acts: 'kill', 'pick up', 'catch'; (e) expressive/communicative actions: 'laugh', 'sing', 'write', 'read'; (f) ethnogeometrical terms: 'edges', 'ends'; (g) life-form words: 'animal', 'bird', 'fish', 'tree'; (h) natural environment: 'the ground', 'the sky', 'the sun', 'water', 'fire', 'day', 'night'; (i) materials: 'wood', 'stone', 'metal', 'glass', 'paper'; (j) mechanical parts: 'wheel', 'pipe', 'wire', 'engine', 'electricity', 'machine'; (k) basic social categories and kin roles: 'men', 'women', 'children', 'mother', 'father'; (l) important cultural concepts: 'money', 'book', 'color', 'number'.

On current evidence it seems likely that some semantic molecules are universal, especially those which are foundational for many other concepts and/or for large lexical classes. 'Hands' is a prime candidate once language-specific polysemy is taken into account, and the same can be argued for 'eyes' (Goddard 2001; Wierzbicka 2007*a*; 2007*b*), for basic social categories like 'men', 'women', and 'children' (Goddard and Wierzbicka to appear), and for the sociobiological concept 'mother', given its foundational status for kinship semantics (Wierzbicka 1992). It is of course clear that many semantic molecules are highly language-specific. In the structure of most complex concepts there are multiple levels of nesting: molecules within molecules. Complex artifact words like *spoon*, *chair*, and *bed*, for example, include physical activity words like 'eat', 'sit', and 'lie' as molecules; they in turn contain body-part concepts, which in turn contain shape descriptors, and they in turn contain the molecule 'hands'.

The concept of semantic molecules appears to have multiple ramifications for our understanding of the overall structuring of the lexicon, for lexical typology, for language acquisition, and for language and cognition studies.

18.3.3 Semantic templates

A semantic template is a structured set of component types shared by words of a particular semantic class—often applicable across many languages. The concept was first employed in explications for artifact and natural kind terms (Wierzbicka 1985), but has recently been elaborated and applied to adjectives (Goddard and Wierzbicka 2007; Wierzbicka 2007*b*; 2008*a*) and to verbs (Goddard and Wierzbicka 2009; Wierzbicka 2009; Wong, Goddard and Wierzbicka to appear). There are affinities with work on lexical templates in other frameworks, e.g., Rappaport Hovav and Levin (1998); Mairal Usón and Faber (2002). Semantic templates vary greatly across semantic domain and word-class. To see this, it is useful to compare templates from two very different domains and word-classes: natural kind terms and physical activity verbs.

Explications for animal terms follow a semantic template with the following sections: [a] CATEGORY, [b] HABITAT, [c] SIZE, [d] BODY, [e] BEHAVIOR, [f] SOUND, [g] RELATION TO PEOPLE. The following is a partial explication—sections [a]–[d] only—for *cats*. The (a) component establishes *cats* as 'animals [M] of one kind'. The (b) components claim that *cats* are conceptualized primarily as domestic animals. The size component (c) is defined in relation to the human body, a kind of anthropocentrism which recurs in countless words of diverse types. The components in (d) identify the distinctive physical features of *cats* as soft fur, a round head with pointy ears, a special kind of eyes, whiskers, a long tail, and soft feet with small sharp claws.

[H] *cats* =>

a.	animals [M] of one kind	CATEGORY
b.	animals [M] of this kind can live with people sometimes they live in places where people live sometimes they live near places where people live	HABITAT
c.	they are not big someone can pick up [M] one with two hands [M]	SIZE

d. they have soft [M] fur [M] BODY
they have a round [M] head [M]
they have pointed [M] ears [M]
their ears [M] are on two sides of the top [M] part of the head [M]
their eyes [M] are not like people's eyes [M]
they have some long [M] hairs [M] near the mouth [M],
they stick out [M] on two sides of the mouth [M]
they have a long [M] tail [M]
they have soft [M] feet [M]
they have small sharp [M] claws [M]

It is important to point out that even the descriptive components of an explication such as [H] are not intended as an external, objective description of the referents. Rather they are aimed at "capturing what is psychologically real and linguistically relevant (from the point of view of native speakers of English)" (Wierzbicka 1996*a*: 344). Terms for most natural kinds and artifacts encapsulate tremendous amounts of cultural knowledge. For example, for *cats* a full explication will include that they chase, catch, and eat small creatures, that they can climb well and move quietly, that they can see in the dark, and that they often sleep for short periods in the day. For animal species with which people have close relationships, such as *cat* (*dog*, *horse*, or *mouse*), the 'behavior' and 'relation with people' sections can run to 10–20 lines of semantic text. Again, these components are not encyclopedic in the sense of representing objective facts about the class of referents. Rather they represent general folk knowledge, encoded in the meaning of the word itself, and in its web of associated phraseology and endonyms (cf. Fillmore's "frame" concept).

> Basic level concepts, including biological ones, are indeed information-rich bundles of perceptual and functional (or better, cultural) features, the two kinds of features being inextricably bound together . . . In fact, the whole folk-generic concept usually has an internal logic to it, so that most, if not all, of its components hang together, and explain and complement one another. (Wierzbicka 1985: 177–8)

We look next at some recent work on the semantic template for verbs. In recent studies, NSM researchers have developed proposals for the structure of several subclasses of physical verbs, including (a) routine bodily processes, like *eating* and *drinking*, and verbs of locomotion, such as *walking*, *running*, *jumping*, (b) verbs of physical contact, such as *hit*, *punch*, *slap*, *kick*, and (c) complex physical activity verbs (typically involving instruments), such as *cutting* and *chopping* (Wierzbicka 2009; Wong, Goddard and Wierzbicka to appear; Sibly 2008; Goddard and Wierzbicka 2009). The overall template structures are very similar, as shown in Table 18.3.

Lexico-syntactic frame refers to the topmost component, with different macro-classes having different frames. For example, intransitive verbs of bodily motion like

Table 18.3. Semantic templates for physical activity verbs of three subclasses

• LEXICO-SYNTACTIC FRAME		
• PROTOTYPICAL MOTIVATIONAL SCENARIO		
routine physical activities, e.g., *eating, drinking, walking, running* • MANNER	physical contact, e.g., *hit, kick, kiss* • HOW THE PERSON USES THE BODY PART	complex physical activities, e.g., *cutting, chopping, grinding, digging* • INSTRUMENT • HOW THE PERSON USES THE INSTRUMENT • WHAT IS HAPPENING TO THE OBJECT
• POTENTIAL OUTCOME		

walking and *running* have the lexico-syntactic frame in (a) below, while complex physical activity verbs like *chopping*, *cutting*, etc. have the frame in (b). The details in the frame determine the mapping from lexical semantics to morphosyntactic expression. The frames define core argument structure, inherent aspect, causal notions, and the controlled nature of the activities.

(a) someone X was doing something somewhere for some time
because of this, this someone's body was moving at the same time in this place, as this someone wanted
(b) someone X was doing something to something Y with something Z for some time
because of this, something was happening at the same time to thing Y, as this someone wanted

A notable feature of these frames is that they are phrased in the imperfective. Most treatments in other frameworks assume without discussion that perfective uses (*walked*, *ran*, *cut*, *chopped*, etc.) are basic, but NSM analysts agree with the tradition in Russian lexicology that, for physical activity verbs, the imperfective forms and uses are semantically simpler than their perfective counterparts. Perfective uses involve extra semantic components, such as the specification that the potential outcome has been achieved. Though we cannot go through the details here, the claim is that this analytical strategy enables a solution to the so-called imperfective paradox and to the problem of how to specify the semantic relationships between constructional variants (syntactic alternations) of a single verb (Levin and Rappaport Hovav 2005).

A distinctive claim of NSM research is that speakers conceptualize human activities by reference to their prototypical motivations. For example, the prototypical motivational scenario for English *walking* states that to say that *someone is walking* is to say that this person is doing as people do when they do something with their legs, etc. because they want 'to be somewhere else after some time'. This does not imply that people only ever *walk* with this motivation; obviously, one can walk for exercise or pleasure, or for other reasons. The claim is that the concept of *walking* makes reference to this particular motivation. Prototypical motivation components can differ considerably in complexity. Complex physical activity verbs (such as *chopping*, *grinding*, *kneading*) have a richer cognitive structure than routine activities, because the former involve a prototypical actor forming a "preparatory thought" directed toward changing the current state of some object. For example, for English *chopping*:

people do something like this when they do something to something hard [M] for some time because a short time before they thought about this something like this:
"I want this something not to be one thing anymore, I want it to be many small things"

Given the goal-directed nature of human action, it is natural that many aspects of the meanings of individual verbs are linked to their prototypical motivation. For example, the prototypical motivation for *chopping* has implications for the kind of instrument needed (something with a sharp edge) and for the manner in which it is used (repeatedly). The NSM approach is unusual in drawing attention to the "intentional" aspects of physical activity verbs, which are sometimes linked with cultural practices and preoccupations. In other approaches such verbs are typically characterized solely in terms of the external, behavioral aspects of situations (e.g., Majid and Bowerman 2007).

For reasons of space, I can illustrate with a full explication for only a single example: *eating* (Wierzbicka 2009). It is important to recognize that, although *eating* is a pretty basic verb in the English lexicon, it is far from being a lexical universal. Languages differ considerably in the precise semantics of verbs for concepts akin to eating, drinking, and so on. Some languages cover both with a single general verb, e.g., Kalam *ñb* 'eat/drink'.

[I] *Someone X was eating something Y*:

LEXICO-SYNTACTIC FRAME

a. someone X was doing something to something Y for some time
because of this, something was happening to this thing Y at the same time

PROTOTYPICAL MOTIVATIONAL SCENARIO

b. people do something like this to something for some time
    if this something is something not like water [M]
when they do something to this something with their mouth [M]
because they want this something to be inside their body

c. when someone does something like this to something for some time, MANNER
    the same thing happens many times
  it happens like this:
    this someone does something to this something with their hands [M]
    at the same time this someone does something to it with their mouth [M]
    because of this, after this, part of this thing is for a short time inside this
      someone's mouth [M]
    when this part is inside this someone's mouth [M], this someone does
      something to it with some parts of their mouth [M]
    because of this, something happens to it at this time
    after this, this someone does something else to it with their mouth [M]
    because of this, after this, it is not inside this someone's mouth [M] anymore
    it is inside another part of this someone's body for some time
POTENTIAL OUTCOME
d. if someone does something like this to something for some time,
    after some time, all parts of this something can be inside this someone's body

In order to illustrate the basic techniques, the examples in this section have been fairly simple words from English. Much NSM work deals with more complex and deeply culturally embedded words, in many languages; in addition to references already cited, see Wierzbicka (1992; 1997; in press), Gladkova (2007*b*), Hasada (2008), Ye (2004; 2007), Bromhead (2009).

18.4 NSM APPROACH TO GRAMMATICAL SEMANTICS

18.4.1 Grammar of semantic primes

As mentioned, semantic primes have an inherent grammar—a "conceptual grammar"—which is the same in all languages. Or to put it another way, each semantic prime has certain combinatorial properties by virtue of the particular concept it represents (Goddard and Wierzbicka 2002*a*: 41–85). The formal realizations (marking patterns, word order, constituent structure, etc.) may differ from language to language without these underlying combinatorial properties being disturbed. The syntactic properties of semantic primes are literally universals of syntax. They can be seen as falling into three kinds: (i) basic combinatorics: for example, that substantive primes and relational substantives can combine with specifiers to form semantic units: THIS THING, SOMEONE ELSE, THE SAME PLACE,

Table 18.4. Some morphosyntactic construction types and associated semantic primes

ONE, TWO, SOME, MUCH/MANY	number-marking systems (incl. duals, paucals)
THE SAME, OTHER	switch-reference, obviation, reflexives, reciprocals
WANT	imperatives, purposives, "uncontrolled" marking
KNOW, SEE, HEAR, SAY	evidential systems
WORDS	delocutive verbs, logophoricity, proper nouns
DO, HAPPEN	case marking and transitivity, passive voice, inchoatives
FEEL, THINK	experiencer constructions, interjections
GOOD, BAD	benefactives, adversatives
BIG, SMALL	diminutives, augmentatives
VERY	superlatives, expressives
NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT	tense (incl. degrees of remoteness), aspect
HERE, ABOVE, BELOW, SIDE, NEAR, FAR	elaborate locational deixis
PART	inalienable possession
KIND	classifier constructions

ONE PART, MANY KINDS, and so on; (ii) an account of basic and extended valencies (see below); (iii) the propositional complement possibilities of primes like KNOW, THINK, and WANT; for example, that KNOW can occur in frames such as I KNOW THAT SOMETHING HAPPENED IN THIS PLACE or I WANT SOMETHING TO HAPPEN NOW. Primes vary widely in the number of alternative valency frames and complementation options available to them. Although groups of primes share particular properties and can be regarded as falling into natural classes, it is equally true that virtually every prime has some idiosyncratic properties, giving each prime a distinctive syntactic signature.

Typological research in the NSM framework indicates that the full set of semantic primes is necessary to capture the semantic content of language-specific grammatical categories and constructions in the world's languages. See Table 18.4 for summary details. This finding runs counter to the claim advanced by some authors that only a subset of the conceptual primes implied in the lexicon are needed for grammatical purposes, or even that the semantic fundamentals of lexical and grammatical semantics are disjoint (Talmy 1988).

In addition to their minimal frames, predicate primes typically allow extended frames in which additional arguments—termed "valency options"—identify or fill out aspects of the situation implied by the predicate. For example, HAPPEN allows us to speak not only of 'something happening' but also of 'something happening to someone' or 'something happening to something'. Borrowing from the usual set of semantic role labels, this additional argument can be labeled an "undergoer" valency option. Likewise, with DO it is possible to add an argument and speak of 'doing something to someone' or 'doing something to something', and the additional argument in this case is conveniently referred to as a "patient" option. This frame can be further extended to speak of 'doing something to something with something', and the additional argument can be labeled as "instrument". Another option for DO is the comitative option.

a.	someone DOES something	[agent]
b1.	someone DOES something to something	[patient$_1$]
b2.	someone DOES something to someone else	[patient$_2$]
c.	someone DOES something to someone/something with something	[instrument]
d.	someone DOES something with someone	[comitative]

Many linguists accept notions of undergoer, agent, patient, and instrument, but typically they are thought of as independent entities of some kind (semantic or thematic roles), rather than as argument slots of basic predicates such as HAPPEN and DO. From an NSM point of view, the idea of an instrument, for example, exists only insofar as one can think about DOING something: it is a conceptual possibility that is "opened up" by the nature of DOING itself, and which is implicitly tied to the concept of DOING. (Jackendoff (1990*a*: 127) explicitly identifies semantic roles as the argument slots of basic predicates, but his basic predicates are abstract conceptual functions, such as AFF "affect", not ordinary word-meanings.)

In relation to patient, the NSM metalanguage forces us to be a little more explicit than a typical definition, such as that offered by Andrews (2007: 137): "a participant which the verb describes as having something happen to it, and as being affected by what happens to it". The technical term 'participant' glosses over the conceptual distinction between persons and things, but when someone *smashes* a plate, for example, we can only describe it in ordinary language as 'someone smashed something', whereas if someone *kills* another person this can only be described as 'someone killed someone else'. This explains why two separate patient frames are shown in (b1) and (b2) above. Recognizing this distinction leads to improved descriptive accuracy. For example, Andrews' characterization of patient could refer to a man who killed *himself*, as well as to a person who killed *someone else*, but in fact most languages distinguish sharply between the two possibilities, treating reflexive sentences as intransitive rather than transitive. Furthermore, many languages have different case marking patterns for the two types of patient (see, for example, Moravcsik 1978; Naess 2007). For further discussion, and discussion of "degrees of transitivity", see Wierzbicka (2002*a*).

In some cases, NSM researchers propose valency options which are seldom recognized in mainstream grammars and which may have no standard labels. For example, it is claimed that semantic prime THINK universally allows a "cognitive topic" option, such that one can say, in all languages, the semantic equivalent of a sentence like 'I was thinking about this someone (this thing, this place, etc.)'. The full valency array for THINK is shown below. Notice that the third and fourth frames show sentential complements: ways in which an expression analogous to a full sentence can be embedded inside the scope of THINK. In many languages the propositional frame is rather restricted in its distribution and range of application (Goddard 2003*a*).

a. someone THINKS about someone/something [cognitive topic]
b. someone THINKS something (good/bad) about
someone/something [topic + complement]
c. someone THINKS like this: "– –" [quasi-quotational thought]
d. (at this time) someone THINKS that $[——]_S$ [propositional complement]

More details about the syntax of predicate primes can be found in Wierzbicka (1996*a*), Goddard and Wierzbicka (2002*b*), Goddard (2008).

18.4.2 Anchoring typological categories in semantic prototypes

That typological comparison rests ultimately on semantic judgments has long been recognized. As Greenberg (1966a: 74) put it: "variation in structure makes it difficult if not impossible to use structural criteria, or only structural criteria, to identify grammatical categories across languages." Greenberg did not shrink from admitting that to identify different category types across languages, in order to compare them, one must rely essentially on semantic criteria. One may also appeal to functional criteria, but on closer inspection functional criteria also depend on semantic judgments, and the same applies to efforts to base cross-linguistic comparison on inventories of situation types, basic domains, conceptual spaces, or whatever. Many grammarians and typologists would agree with Wierzbicka's (1998) summary:

> ... the grammatical resources of any language are limited. Often, therefore, a grammatical construction is centred around a prototypical meaning, and has also various extended uses, accommodating other meanings (usually related to the prototype). Often, the same prototypical meanings recur in different languages, whereas the extensions are language-specific. (Wierzbicka 1998: 143)

In most typological work, the details of presumed prototypes are stated in complex, English-specific terms. To give a clearer idea of how complex categories can be treated from an NSM prototype-theoretic point of view, we will look at two examples—one from morphology (cases) and one from syntax (relative clauses).

Cases. As early as her 1980 *The Case for Surface Case*, Wierzbicka (1980) was arguing that inflectional cases are best dealt with by way of a prototype-with-extensions analysis (rather than, for example, with an abstract "general meaning"

(*Gesammtbedeutung*) along the lines proposed by Roman Jakobson in his celebrated analysis of Slavic cases). In recent work focused on Polish (Wierzbicka 2008*b*), she reiterates her claim that each of the four cases traditionally labeled nominative, accusative, instrumental, and dative has a semantic prototype concerned with a scenario of human action.

In one prototypical scenario, the speaker is talking simply about someone doing something. In a second scenario, the speaker is talking about someone who did something to something and who wanted something to happen to this thing. The doer is still marked by the nominative, and the target object by the accusative. If the speaker is talking about someone who was doing something to something with some other thing, the "other thing" is marked by the instrumental. In a fourth prototypical scenario, the speaker is talking about someone who did something to something because he or she wanted something to happen to someone else, in which case the affected person is marked with the dative. The lexical semantics of the verb 'give' make it a natural candidate for a dative-marked recipient. Using NSM, one can formulate these four prototypical scenarios as follows:

Scenario I: a semantic prototype for the nominative

<u>someone</u> is doing something

Scenario II: a semantic prototype for the accusative

someone did something to <u>something</u>
  because this someone wanted something to happen to this thing
something happened to this thing because of it

Scenario III: a semantic prototype for the instrumental

someone was doing something to something with <u>something else</u> for some time
  because this someone wanted something to happen to this thing

Scenario IV: a semantic prototype for the dative

someone did something to something
  because this someone wanted something to happen to <u>someone else</u>
something happened to this other someone because of it

These scenarios can be used as stable conceptual reference points in deciding whether a particular noun marker in a given language warrants being identified as a nominative, an accusative, an instrumental, or a dative.[2] The simple and language-independent wording is much preferable to perennially contested, and

[2] The scenarios differ in tense and aspect in the interests of psychological and linguistic plausibility. For example, the past tense of Scenario II (accusative) is connected with the speaker's interest in the result on the object, implied by the final component; the durative component in Scenario III (instrumental) is connected with speaker's likely interest in the process, implied by the mention of the instrument. For a full explanation, see Wierzbicka (2008*b*).

English-specific, terms such as "instigator", "affected", "volitional", "experiencer", "beneficiary", and the like.

In any given language a particular case will be used not only in its prototypical context, but also in a set of extended uses with interrelated semantics. For example, Polish nominative–dative constructions with a verb of 'doing' can be divided into five major classes (Wierzbicka 2008*b*). All five share the idea of someone doing something, wanting something to happen to someone else, but they differ in the nature of the "potential effect", i.e., whether because of the subject's action the dative-marked referent (i) 'can have something', (ii) 'can do something', (iii) 'can see/hear/know something', (iv) 'can feel something good/bad', or (v) 'can feel something in his/her body'. It is possible here to mention only a couple of the many subtleties captured and explained by these analyses. First, the construal under (i) allows not only for transfer of ownership (as in *sprzedać* 'sell') or physical transfer (as in *rzucić* 'throw') but also for 'buying', 'making', or 'sewing' something for someone, and the like. Second, because it mentions only potential 'having', it does not absolutely guarantee that the other person did have something as a result (for example, one can *send* something and it can go astray). Nonetheless, and this is the third point, the potential must be real, intention alone is not enough. This explains a contrast with the near-paraphrase with preposition *dla* 'for'. For example, sentence (1a) with *dla* 'for' can be expanded with 'not knowing that he had died'; but the dative sentence in (1b) cannot be expanded in this way. The construal under (ii) might seem identical to that implied by the English "internal dative", but English sentences like *Peter opened Paul a tin of sardines* always imply a tangible effect on the object, whereas the Polish construction only requires that it be clear what the person designated by the dative can do as a result. Hence a sentence like (2), about opening a door for someone, is fine in Polish with the dative, though it can scarcely be translated with an English internal dative.

(1) a. *Kupiła* *dla* *niego* *sweater.*
buy.PAST.3SG.FEM for he.GEN sweater.ACC
'She bought a sweater for him.'

b. *Kupiła* *mu* *sweater.*
buy.PAST.3SG.FEM he.DAT sweater.ACC
'She bought him a sweater.'

(2) *Piotr* *otworzył* *Pawłowi* *drzwi.*
Peter.NOM open.PAST.3SG.MASC Paul.DAT door.ACC
'Peter opened the door for Paul.'

Technical descriptors like "beneficiary", "experiencer", or "affected" do not provide enough clarity or simplicity to match the predictiveness of the NSM explications.

Relative clause. The most influential functional definition for a relative clause is that of Keenan and Comrie (1977: 63–4). They say that any "syntactic Object" is a relative clause if "it specifies a set of objects . . . in two steps: a larger set is specified, called the domain of relativisation, and then restricted to some subset of which a certain sentence is true. The domain of relativisation is expressed in the surface structure by the head NP, and the restricting sentence by the restricting clause".

Despite its canonical status in discussions of relative clauses, this characterization is highly implausible from a cognitive point of view (Wierzbicka 1998). For example, for a sentence such as the following (adduced by Keenan 1985: 142): *I picked up two towels that were lying on the floor*, it would hardly be plausible to suppose that the speaker has in mind the set of possible pairs of towels and that the function of the relative clause consists in narrowing this set down to just one such pair. Rather, what the speaker appears to be doing is providing some additional information about the two towels referred to in the main clause. Wierzbicka proposes the following schema:

[J] *I picked up two towels that were lying on the floor*:

I say: I picked up two towels
I want to say something more about these two towels at the same time:
  [I say] they were lying on the floor

Pursuing her critique of the "subset" characterization further, Wierzbicka (1998: 186) adduces the following sentences (from a contemporary novel): (a) *Snow that was drowning the city . . .* (b) *How could he trust even this circle of elastic on the sleeve of the girl's frock that gripped her arm?* In relation to (a) she comments: "It seems really beyond belief that the speaker is thinking here about the set of all snows and delimiting the subset of snow that was drowning the city". In relation to (b), it is not very plausible that the function of the relative clause is one of identification, since the preceding phrase "on the sleeve of the girl's frock" would seem to provide adequate identification: "Rather, the clause *that gripped her arm* provides additional information about the elastic in question—information that the speaker sees as relevant to the content of the main clause and wants to integrate with it."

The semantic formula 'I want to say something more about this thing at the same time' seems to capture the intended meaning of these relative clauses adequately.[3] Wierzbicka goes further to suggest that it constitutes a clear and appropriate characterization of the prototypical concept of a relative clause. (This proposal assumes that prototypical relative clauses are unspecified with respect to the distinction between restrictive and non-restrictive ones—a distinction which very few languages seem to draw in any consistent way, and which is often vague in English; cf. Fox and Thompson (1990).)

[3] It is not being claimed that the formula in [J] applies to all English relative clauses. For example, it does not exactly fit relative clauses with indefinite or generic NP heads, e.g., *It's the only place that carries this book*; also, there are some relative clauses which, in combination with a determiner, do seem to indicate a subset reading, e.g., *Those who went east found water and survived.*

18.4.3 Grammatical polysemy and ethnosyntax

Two NSM concepts in grammatical semantics which cannot be fully illustrated here for reasons of space are grammatical polysemy and ethnosyntax. The idea behind grammatical polysemy is simply that grammatical constructions may exhibit polysemy which can be independent (to some extent) of the lexical items involved in the construction. The Polish dative, outlined above, is one example. Some other examples explored at length in the NSM literature include Wierzbicka (1988) on the English *have a VP-INF* construction (*have a chat*, *have a look*, etc.) and the Japanese adversative passive, Wierzbicka (2002*b*) on English *let*-constructions, and Goddard (2003*c*) on the Malay dynamic *ter*- prefix. Given the date of the earliest of these studies, Wierzbicka deserves to be seen as one of the precursors of construction grammar.

The term "ethnosyntax" (Wierzbicka 1979) refers to inquiry into phenomena at the intersection of grammar, semantics, and culture. Wierzbicka argued that the natural semantic metalanguage promised to bring new rigor into an area of study anticipated by von Humboldt, Bally, Baudoin de Courtenay, Boas, Sapir, and Whorf; namely, the study of the "philosophies" (or ethnophilosophies) built into the grammar of different languages. Examples of NSM studies in ethnosyntax include Wierzbicka's (1992) studies of fatalism in Russian dative-subject constructions (cf. Goddard 2002; 2003*b*), Travis's explorations of the semantics of the diminutive and ethical dative in Spanish (2004; 2006),Wong (2004) on nominal reduplication in Singapore English, and Priestley (2008) on inalienable possession in Koromu (Papua New Guinea). Of particular interest, given the rise of English as a global lingua franca, are studies of the ethnosyntax of English; for example, English epistemic adverbs (Wierzbicka 2006*a*), wh-imperatives (Wierzbicka 2003 [1991]), and tag questions (Wong 2008).

18.5 Cultural scripts and ethnopragmatics

The NSM approach has a "sister theory" in the form of the theory of cultural scripts. Studies of communicative style usually assume that in any particular speech community there are certain shared understandings (norms of interpretation, rules of speaking, discourse strategies, etc.) about how it is appropriate to speak in particular, culturally construed, situations. How can such norms be stated in a clear, testable, and non-ethnocentric fashion? Conventional labels such as "directness", "formality", "involvement", "politeness", etc. are useful up to a point but are somewhat vague and shifting in their meanings, in the sense that they are used with different meanings by different authors and in different contexts. Furthermore, such terms bring with them an element of ethnocentrism (specifically, Anglocentrism), because the relevant concepts are not usually found in the cultures being described and cannot be translated easily into the languages involved. The NSM solution is to formulate hypotheses about culture-specific norms of communication using the metalanguage of universal semantic primes. A cultural norm formulated in this way is referred to as a "cultural script" (Wierzbicka 2003 [1991]; 1996*b*; Goddard and Wierzbicka 1997; 2004).

To see what cultural scripts look like, we will consider two cultural scripts proposed in the *Ethnopragmatics* collection (Goddard 2006*b*). Their most notable feature is their intelligibility and the simplicity of the phrasing of individual phrases and sentences, but taken as a whole each script captures a highly specific and quite complex configuration. Script [K] comes from Ye's (2006: 152–3) study of the semiotics and associated cultural norms of Chinese facial expressions. The script captures a social proscription against allowing others to detect in one's face any sign that one is feeling 'something very good' or 'something very bad' on account of some personal good fortune or ill fortune.

[K] *A Chinese cultural script for concealing displays of 'feeling very good/bad'*

[people think like this:]
when someone feels something very good/bad because something very good/bad happens to this someone,
it is not good if other people can know this when they see this someone's *liǎn* ('face') [M]

The script in [L] is proposed (Wong 2006: 116) to capture a Chinese Singaporean attitude (no doubt widespread across the "Sinosphere") which underlies the use of honorific kin terms such as *Auntie* and *Uncle*. The first part of the script indicates that people are, so to speak, "tuned" to thinking of other people in terms of relative age. The second part prescribes a certain attitude toward such people (roughly, thinking of them as different from oneself and as "above" oneself) and also mandates some positive views about them.

[L] *A Singapore English cultural script for "respectful" attitude toward someone older*

[people think like this:]
I can think about some other people like this:
"I have lived for some time, these people have lived for some time more"
if I think like this about someone, I have to think about this someone like this because of it:
"this someone is not someone like me, this someone is someone above me"
I have to think something good about this someone because of this

This script is only one of a suite of age-related Chinese cultural scripts, some of which articulate more specific attitudes linked with generational differences. These scripts enable much more detail than the normal simplistic description in terms of "respect for age", which glosses over important differences, between, for example, Chinese and Korean norms (Yoon 2004) in regard to age.

It is important to stress that despite the possible connotations of the word 'script', cultural scripts are not binding on individuals. They are not proposed as rules of behavior, or as descriptions of behavior, but as normative rules of interpretation and evaluation. It is up to individuals in concrete situations whether to follow (or appear to follow) culturally endorsed principles and, if so, to what extent; or whether to defy, manipulate, or subvert them, play creatively with them, etc. Whether or not cultural scripts are being followed in behavioral terms, however, the claim is that they constitute a kind of shared interpretive background. It also has to be stressed that a few simple examples cannot give an accurate impression of the complex inter-relationships between and among the large number of scripts operative in any culture, including various forms of intertextuality, e.g., some being more general than others, some taking priority over others, some competing with others. Equally, it is clear that many scripts must be tailored to particular types of interlocutors, settings, and discourse genres.

One of the key concerns of much work in the cultural scripts framework is to "de-naturalize" the pragmatics of English, which is often taken (or, mistaken) as culturally unmarked; cf. Wierzbicka (2003 [1991]; 2006*a*), Peeters (2000); Goddard (2006*a*). It therefore seems important to adduce at least one cultural script of mainstream Anglo culture. In doing so we also take the opportunity to show how ethnopragmatics is not solely a matter of usage conventions but can exert an influence on language structure. A wide range of sociological, historical, and culture-analytical literature indicates that something like "personal autonomy" is one of the primary ideals of Anglo culture. Script [M] is intended to capture an important aspect of this ideal.

[M] *Anglo cultural script for "personal autonomy"*

[people think like this:]
when someone does something, it is good if this someone can think about it like this:
"I am doing this because I want to do it"

It is not difficult to see that this ideal can inhibit speakers of mainstream English from using the bare imperative when they want someone to do something (because a bare imperative includes a message like: 'I want you to do this; I think that you will do it because of this'). It is well known that in most social situations Anglo speakers prefer to frame their directives in a more elaborated fashion, using "interrogative directives" (wh-imperatives) such as *Will you ...?*, *Would you ...?*, *Can you ...?*, *Could you ...?*, *Would you mind ...?*, and the like. Although these constructions clearly convey the message 'I want you to do this', they acknowledge the addressee's autonomy by embedding the potentially confronting message into a question form, as if inviting the addressee to say whether or not he or she will comply. Another favored strategy is the use of "helpful suggestions", such as *Perhaps you could*..., *You might like to*..., and *I would suggest*... (Wierzbicka 2006*b*: 51f).

In a similar fashion, Wierzbicka and others have argued that Anglo cultural values encourage speakers to express something like epistemic reserve when saying what they think, and to routinely acknowledge the possible existence of differing opinions (Wierzbicka 2003 [1991]; 2006*a*). This, it is argued, is linked with the high-frequency English formula *I think*, phrases like *in my opinion* and *as I see it*, hedges such as *kind of* and *a bit*, and also with the frequency and grammatical elaboration of tag questions in mainstream Anglo English (Wong 2008).

Explaining these ways of speaking in terms of culture-specific Anglo values, such as personal autonomy, is quite different to attributing them to "universals of politeness", in the style of Brown and Levinson (1987) and later versions of neo-Griceanism. Wierzbicka has long been a strong critic of neo-Gricean pragmatics, charging it with semantic naivety, explanatory inadequacy, and thinly disguised Anglocentrism (evident both in its individualist orientation and in its key terms, such as "imposition").

Culture-specific pragmatic norms tend to spawn semantically specialized constructions which are tailor-made to meet the communicative priorities of the culture, as routinized patterns of usage "harden" into fixed morphosyntactic constructions (Traugott and König 1991; Evans and Wilkins 2000: 580–5). Because it uses the same metalanguage to depict meaning in both semantics and pragmatics, the NSM approach allows for a particularly clear account of how the semanticization of pragmatic implicature works as a process of language change.

18.6 Concluding remark

The NSM approach offers a comprehensive and versatile approach to meaning analysis: highly constrained and systematic, non-ethnocentric, and capable of producing representations with high cognitive plausibility. Given the pervasiveness of meaning-based and meaning-related phenomena in languages (in lexicon, morphology, syntax, prosody, and pragmatics), the approach surely has a tremendous amount to offer linguistics at large. Of course, NSM is not a complete theory or methodology of linguistic analysis. If languages can be thought of as systems for correlating meanings with forms, NSM's strengths lie on the meaning side of the equation. There can be little argument, however, that the linguistics of the twentieth century concentrated predominantly on form, at the expense of meaning, and that it was the poorer because of it. Hopefully the twenty-first century will see the balance restored, so that meaning can re-assume a central place in linguistics.

Appendix A

Semantic primes in two additional languages—Japanese (Hasada 2008) and Russian (Gladkova 2007*a*)

watashi *I*, anata *you*, dareka *someone*, nanika/mono/koto *something/thing*, hito/hitobito *people*, karada *body*	ja *I*, ty *you*, kto-to *someone*, čto-to/vešč' *something/thing*, ljudi *people*, telo *body*
shurui *kind*, bubun *part*	rod/vid *kind*, čast' *part*
kore *this*, onaji *the same*, hoka *other*	ėtot *this*, tot že *the same*, drugoj *other*
hito-/ichi- *one*, futa-/ni- *two*, ikutsuka *some*, minna *all*, takusan *much/many*	odin *one*, dva *two*, nekotorye *some*, vse *all*, mnogo *much/many*
ii *good*, warui *bad*	xorošij/xorošo *good*, ploxoj/ploxo *bad*
ookii *big*, chiisai *small*	bol'šoj *big*, malen'kij *small*
omou *think*, shiru *know*, hoshii/-tai/nozomu *want*, kanjiru *feel*, miru *see*, kiku *hear*	dumat' *think*, znat' *know*, xotet' *want*, čuvstvovat' *feel*, videt' *see*, slyšat' *hear*
iu *say*, kotoba *words*, hontoo *true*	govorit'/skazat' *say*, slova *words*, pravda *true*
suru *do*, okoru/okiru *happen*, ugoku *move*, fureru *touch*	delat' *do*, proisxodit'/slučat'sja *happen*, dvigat'sja *move*, kasat'sja *touch*
(dokoka) iru/aru *be (somewhere)*, iru/aru *there is*, motsu *have*, (dareka/nanika) dearu *be (someone/something)*	byt' (gde-to) *be (somewhere)*, byt'/est' *there is*, byt' u *have*, byt' (kem-to/čem-to) *be (someone/something)*
ikiru *live*, shinu *die*	žit' *live*, umeret' *die*
itsu/toki *when/time*, ima *now*, mae *before*, ato *after*, nagai aida *a long time*, mijikai aida *a short time*, shibaraku no aida *for some time*, suguni *moment*	kogda/vremja *when/time*, sejčas *now*, do *before*, posle *after*, dolgo *a long time*, korotkoe vremja *a short time*, nekotoroe vremja *for some time*, moment *moment*
doko/tokoro *where/place*, koko *here*, ue *above*, shita *below*, tooi *far*, chikai *near*, men *side*, naka *inside*	gde/mesto *where/place*, zdes' *here*, nad *above*, pod *below*, daleko *far*, blizko *near*, storona *side*, vnutri *inside*
-nai *not*, tabun *maybe*, dekiru *can*, -kara *because*, moshi (ba) *if*	ne *not*, možet byt' *maybe*, moč' *can*, potomu čto *because*, esli *if*
sugoku *very*, motto *more*	očen' *very*, bol'še/ešče *more*
yoo/doo/yooni *like/how/as*	kak/tak *like*

CHAPTER 19

LINGUISTIC MINIMALISM

CEDRIC BOECKX

Linguistic minimalism refers to a family of approaches exploring a conjecture, first formulated by Noam Chomsky in the early 1990s, concerning the nature of the human language faculty. My aim in this chapter is fourfold. First, I want to state as clearly as I can what the conjecture amounts to, what sort of research program emerges from it, and how it could be carried out (using examples from the existing literature as concrete illustrations). Second, I want to emphasize that the minimalist program for linguistic theory did not arise out of nowhere. It is firmly grounded in the generative enterprise and the rationalist ("Cartesian") tradition more generally. Third, the pursuit of specific minimalist analyses follows a certain research style, often called the "Galilean style", whose core properties I want to discuss, since they help one understand why certain moves are made when minimalism is put into practice. Fourth, I want to highlight the fact that minimalism, if rigorously pursued, naturally gives rise to a specific way of approaching interdisciplinary problems such as "Darwin's Problem" (the logical problem of language evolution). Indeed, I believe that minimalism has significantly contributed to the resurgence of "biolinguistic" themes in recent years, and may mark a return of linguistic studies to

As I stress at the outset of this chapter, this overview is very much minimalism as I see it. But I am indebted to a number of people who opened my eyes to both big and small points over the years, and changed my vista several times along the way: Noam Chomsky, Juan Uriagereka, Norbert Hornstein, Massimo Piattelli-Palmarini, Paul Pietroski, Marc Hauser, Sam Epstein, Bob Berwick, Howard Lasnik, Željko Bošković, Dennis Ott, Bridget Samuels, Hiroki Narita, and Ángel Gallego.

the heart of cognitive science after two decades of (unavoidable, indeed, necessary) modularization.

I should stress right from the start that the overview that follows is obviously a very personal one. The reader should bear in mind that, although I rely on the works of a large group of researchers, this is very much "linguistic minimalism as I see it".

19.1 Beyond explanatory adequacy

Science is all about understanding, not describing. Scientists are in the business of explaining, not merely cataloging. Since its inception, the generative approach to language has studied language as a natural object, pretty much the way a physicist or chemist approaches his object of study. Since Chomsky's early works, language (more accurately, the language faculty [FL]) has been treated as an organ of the mind in order to shed light on one "big fact" about human beings: short of pathology or highly unusual environmental circumstances, they all acquire at least one language by the time they reach puberty (at the latest), in a way that is remarkably uniform and relatively effortless. The acquisition of language is all the more remarkable when we take into account the enormous gap between what human adults (tacitly) know about their language and the evidence that is available to them during the acquisition process. It should be obvious to anyone that the linguistic input a child receives is radically impoverished and extremely fragmentary when compared with the subtlety and complexity of what the child acquires. It is in order to cope with this "poverty of stimulus" that Chomsky claimed that humans are biologically endowed with an ability to grow a language, much like ethologists have done to account for the range of quite specific and elaborate behaviors that animals display. The biological equipment that makes language acquisition possible is called Universal Grammar (UG). Fifty years of intensive research have revealed an astounding array of properties that must be part of UG if we are to describe the sort of processes that manifest themselves in all languages (in often quite subtle ways) and the way such processes become part of the adult state of FL. This much generative grammarians (over a broad spectrum of theoretical persuasions) take as undeniable. Around 1980, the array of UG properties just alluded to crystallized into a framework known as the Principles-and-Parameters (P&P) model, which quickly became the mainstream or standard model for many linguists. The P&P model conceives of UG as a set of principles regulating the shape of all languages and a set of parameters giving rise to the specific forms of individual languages. Principles of UG can be thought of as laws to which all languages must abide. Among other things, these account for why all languages manifest recursive dependencies that go beyond the power of finite-state machines (as in *anti-missile missile, anti-anti-missile-missile missile*, etc.), and for the existence of displacement (situations where an element is interpreted in a very different location from the one it is pronounced in, as in *What did John say that Mary saw* __?). They capture the range of arguments a verb can take to express certain events (*John ate/John ate the food*, but not *John saw/John ate Mary the food*). They also account for the fact that certain dependencies, even "long-distance" ones, must be established within certain local domains (*What did you say she saw* __? but not *What did you ask who saw* __?). And they also require hierarchical structures in syntax (phrases) to be dominated ("headed") by a designated element inside them (a phenomenon known as "endocentricity") (e.g., there is no noun phrase without a noun at the center of it: *[John's happiness]* but not *[John's arrive]*). UG principles also account for why certain elements must be pronounced in some positions in sentences (*John was arrested*), but not in others (*was arrested John*), and so on. The degree of details with which UG principles are formulated requires advanced training in linguistics (see Lasnik and Uriagereka 1988, Haegeman 1994), and never fails to impress or overwhelm the non-specialist. But they have enabled practitioners to account for what looks like fundamental properties of FL.

Next to these invariant principles linguists recognize the existence of parameters to account for the fact that superficially there is more than one language. If all principles were invariant down to their last details, there would only be one language heard/signed on the planet. Since this is not the case, some aspects of UG must be left "open" for the local linguistic environment to act upon. We can think of parameters as items on a (UG) menu; a pre-specified set of options from which to choose. Some languages will have a Subject–Verb–Object word order (English), others will have a Subject–Object–Verb order (Japanese). Some language will require the question word to be at the front of the sentence (English), others not (Chinese). And so on. Thus conceived, the P&P model likens the acquisition task to choosing the right options (setting the right values of the parameters) to conform to the language of the environment, much like one switches the various buttons of a circuit to on or off to achieve the desired configuration. Arguably for the first time in history, the P&P model allowed linguists to resolve the tension between the universal and particular aspects of language. It led to a truly impressive range of detailed results covering a great number of languages (see Baker 2001), and accounted for key findings in the acquisition literature (see Guasti 2002, Wexler 2004, Snyder 2007).

By the end of the 1980s, after more than ten years of sustained effort revealing the fine structure of P&P, Chomsky got the impression that the overall approach was well-established, and that it was time to take the next step on the research agenda of the generative enterprise. The next step amounts to an attempt to go beyond explanatory adequacy. Chomsky 1965 distinguishes between three kinds of adequacies: observational, descriptive, and explanatory, and, not surprisingly, puts a premium on explanatory adequacy. The aim of (generative) linguistics was first and foremost to account for the amazing feat of human language acquisition. Once it was felt that the P&P model met this objective (in some idealized fashion, of course, since no definitive P&P theory exists), it became natural to ask how one can make sense of the properties the P&P model exhibits—how much sense can we make of this architecture of FL? Put differently, why does FL have this sort of architecture?

Quite reasonably, Chomsky formulated this quest beyond explanatory adequacy in the most ambitious form (what is known as the strong minimalist thesis), in the form of a challenge to the linguistic community: Can it be shown that the computational system at the core of FL is optimally or perfectly designed to meet the demands on the systems of the mind/brain it interacts with? By optimal or perfect design Chomsky meant to explore the idea that all properties of the computational system of language can be made to follow from minimal design specifications, a.k.a. "bare output conditions"—the sorts of properties that the system would have to have to be usable at all (e.g., all expressions generated by the computational system should be legible, i.e., formatted in a way that the external systems can handle/work with). Put yet another way, the computational system of language, minimalistically construed, would consist solely of the most efficient algorithm to interface with the other components of the mind, the simplest procedure to compute (generate) its outputs (expressions) and communicate them to the organs of the mind that will interpret them and allow them to enter into thought and action. If the strong minimalist thesis were true, FL would be an ideal linguistic system. But it should be stressed that the point of the minimalist program is not to prove the validity of this extreme thesis but to see how far the thesis can take us, how productive this mode of investigation can be. The strong minimalist thesis amounts to asking whether we can make perfect sense of FL. Asking this question is the best way to find out how much sense we can make out of FL. The points where the minimalist program fails will mark the limits of our understanding. If one cannot make perfect sense of some property P of FL (i.e., if P cannot be given a minimalist rationale in terms of computational efficiency toward interface demands), then P is just something one must live with, some accident of history, a quirk of brain evolution, some aspect of FL that one must recognize in some brute force fashion, one whose secrets must be forever hidden from us, as Hume might have said.

There is no question that the minimalist program is the right strategy to account for properties of FL. Its conceptual/methodological legitimacy can hardly be questioned (except perhaps by appealing to the fact that biological organs in general do not display the sort of optimality that the minimalist program is looking for—the "tinkering" side of biology to which I return in section 19.4 below), but the timing of its formulation may be. A program such as minimalism is neither right nor wrong; it is either fertile or sterile. Its success will depend on the state of maturity reached by the enterprise within which it is formulated. It is quite possible that minimalist questions in linguistics are premature. It very much depends on how one feels about the P&P model. It is important here to stress the term "model". The feasibility of the minimalist program does not rely on the accuracy of all the principles and parameters posited down to their smallest details (indeed such principles and parameters are constantly revised, enriched, improved upon, etc.), but it does depend on whether we think that the sort of approach defined by the P&P model has a fighting chance of being explanatorily adequate. I side with Chomsky in thinking that it does, but, not surprisingly, P&P skeptics have found the minimalist program outrageous.

19.2 A guide to minimalist analysis

Having made its immediate conceptual foundation clear, let me now turn to how the strong minimalist thesis could be put into practice (for a more extended discussion, see Boeckx 2006, ch. 5; see also Boeckx 2008*a*; Lasnik et al. 2005; Hornstein et al. 2006; and Uriagereka 1998).

The first thing for me to note here is something I already alluded to above: the point of minimalist inquiry is not to pick a definition of optimal design and prove its existence but rather to look for one that allows us to make progress at the explanatory level. (As Putnam 1962 observes in the context of Einstein's principle that all physical laws be Lorentz-invariant, it is perhaps because of their vagueness, their programmatic nature, that scientists find such guiding principles extremely useful.) This is another way of saying that there are many ways of articulating minimalist desiderata. There are, in fact, two, possibly three, major families of approaches that presently fall under the rubric of linguistic minimalism. All of them grew out of Chomsky's early minimalist writings (Chomsky 1993, 1995), so I will begin this section by giving the flavor of the early minimalist period before turning to more recent developments.

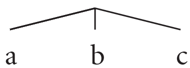
Among the generalizations arrived at in the elaboration of the P&P model was one that proved particularly instrumental in the development of the minimalist program. Chomsky (1986*a*:199) interpreted the unacceptability of sentences like *was believed John to be ill* and *John was believed is ill* (compare *John was believed to be ill)* as indicating that an element had to be displaced at least once, but could not be displaced twice to a case/agreement position (preverbal position triggering subject-agreement on the verb). Put differently, displacement of a noun phrase (out of the domain where it is obviously interpreted, in this case, the vicinity of *be ill*) must happen until that noun phrase reaches a case-assigning/agreement-triggering position. But once it has reached that position, the displaced element is frozen there. From this Chomsky concluded that movement to a preverbal subject position was "a last resort operation". In more general terms, Chomsky claimed that some operation must take place, and, further, that once the operation has taken place, it cannot happen again. Chomsky took this to mean that the application of certain operations is banned if nothing is gained by performing it.

In a similar vein, at the end of the 1980s, Chomsky and others began to reinterpret some generalizations and principles in terms of least effort strategies. Take, for example, the so-called Minimal Distance Principle. Rosenbaum formulated this principle in 1970 to deal with instances of so-called control. Control is a cover term for the mechanism that lies behind the way we interpret sentences like *John tried to leave* as indicating that *John* is both the agent of "trying" and of "leaving". That is, we interpret the sentence *John tried to leave* as meaning that John did something that would make it possible for him (and not somebody else) to leave. Control is also at work in sentences like *John persuaded Mary to leave*. The grammar dictates that this sentence be understood as meaning that the leaver is Mary not John. Rosenbaum's Minimal Distance Principle expresses the idea that the element understood as the subject of the infinitival clause (*to leave* in our examples) is the element that is closest to that infinitival clause. This, too, has the flavor of an economy/least effort condition, and was interpreted as such in a minimalist context by Hornstein (1999).

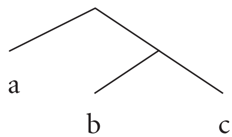
From the early 1990s onward, Least Effort and Last Resort principles became a cornerstone of syntactic theorizing, a key feature of syntactic operations. In addition to claiming that syntax should be organized around principles that legislate against superfluous steps in derivations and superfluous elements in representations, Chomsky also suggested that the architecture of syntax followed from "virtual conceptual necessity".

For example, the fact that sentences are made up of a potentially infinite number of distinct phrases has been taken to force upon linguistic theory a grouping operation which combines at least two elements a and b forming a set $\{a, b\}$. This is the operation Chomsky calls *Merge*. Sharpening the use of virtual conceptual necessity, Chomsky claimed that since *at least* two elements must be the input of Merge, let us say that *at most* two elements must be the input of Merge. This means that if we want to combine three elements into a set (phrase), two applications of Merge are required. A first step that puts two elements together, and a second step that takes the group just formed and joins the third element to it. This captures Kayne's (1984) binary branching requirement on syntactic phrases that had become standard in the P&P model.

(1) BAD:



GOOD:



The idea that this piece of P&P syntax may follow from virtual conceptual necessity has recently been related to various suggestions to the effect that binary branching trees (which, at a suitable level of abstraction, are pervasive in nature) may be computationally more efficient/economical than other kinds of representations (see, for example Medeiros 2008, Boeckx 2008*b*). If this turns out to be the case, this convergence of virtual conceptual necessity and computational economy/efficiency is the type of result that scientists would regard as strongly suggesting that the minimalist program is on the right track.

Chomsky also pointed out that refraining from imposing an upper bound on the number of applications on Merge yields recursive structures, and thus captures the essence of what allows language to make infinite use of finite means. Likewise allowing Merge to recombine members of the sets it forms—what Chomsky recently called internal merge—yields a version of displacement. Note that no additional condition is needed to allow displacement. In fact, as Chomsky (2004) points out, it would take an extra condition to disallow it. Note also the pleasing symmetry between the operation that yields displacement and the basic operations that combines two elements (Merge). The emergence of economy conditions on derivations and representations, the consequences of virtual conceptual necessity, and the search for unity and symmetry in syntactic operations and representations now define what we take to constitute the true character of linguistic principles. Such guidelines play themselves out in different ways, depending on which particular one is stressed. For example, in the wake of Baker's (1985) Mirror Principle and Kayne's (1994) Linear Correspondence Axiom, several researchers have explored the degree of transparency between syntactic representations and morphological or linear orders. In practice this has led to the so-called Cartographic Project, which connects the robust restrictions one observes at the level of morphological make-up of words and the linear order of elements in a sentence to very rich, fine-grained, highly articulated series of phrases in the clause (see Cinque 1999, Rizzi 1997). The approach is animated by minimalist concerns, as it pays attention to the nature of the mapping between narrow syntax and the interfaces, and also as it reflects on the nature of the kind of computation needed and its cost. As Rizzi (2004: 9) puts it "one driving factor of the cartographic endeavor is a fundamental intuition of simplicity (…). Complex structures arise from the proliferation of extremely simple structural units: ideally, one structural unit (a head and the phrase it projects) is defined by a single syntactically relevant feature". Rizzi (2004: 10) goes on to point out that "local simplicity is preserved by natural languages at the price of accepting a higher global complexity, through the proliferation of structural units.... Recursion is cheap; local computation is expensive and to be reduced to the minimum".

A different, though ultimately related, kind of approach capitalizes on Chomsky's 1993 idea that syntactic operations are subject to Last Resort. Such an idea, first formulated in the context of movement, has been extended to all syntactic relations, and has led to the development of highly constrained, so-called crash-proof models of grammar, where lexical features and their mode of licensing plays a significant role (see Adger 2003; Frampton and Gutmann 2002, among others). This line of inquiry connects with the Cartographic Project in that it leads to elaborate feature organizations (geometries) that mirror the series of heads posited in Cartographic studies. I anticipate an important degree of convergence between these two families of approaches in the near future.

A third type of minimalist studies, which emerged more recently, has de-emphasized the role of specific features in driving syntactic computations and paid more attention to the consequences of assuming a more derivational architecture, where small chunks of syntactic trees (aka "phases") are sent cyclically to the interfaces (see Epstein et al. 1998, Uriagereka 1999, Chomsky 2000*a*). This type of approach (articulated by Chomsky over a series of papers beginning with Chomsky 2000*a*; see Chomsky 2001*b*, 2004, 2007, 2008) seeks to turn the economy principles of the early minimalist period into theorems, and generally minimizes the size of the output sent to the interfaces. Because they adhere to an even more minimal inventory of properties of lexical items ("features") and structures, such studies only ensure the bare minimum (legibility requirement) at the interfaces. As a result, a fair amount of filtering must be performed by the external systems to ultimately characterize "well-formed" expressions. In my view, this marks a partial return to the Filtering architecture that characterized much of the P&P era (see Chomsky and Lasnik 1977, Lasnik and Saito 1984, 1992), and that goes back in many ways to Chomsky and Miller (1963). The studies under consideration relate rather naturally to models of the lexicon and morphology that deflate the pre-syntactic lexicon (Hale and Keyser 1993, 2002; Halle and Marantz 1993, Marantz 2000, 2006; Borer 2005). They also fit particularly well with Neo-Davidsonian semantic representations (Pietroski 2003, Boeckx 2008*b*).

It is impossible for me to review in a brief overview like this one the range of results that minimalist theorizing has already achieved, but I would like to highlight three conclusions that seem to be gaining significance and plausibility as minimalist inquiry progresses. The first pertains to the external systems to which the core computational system relates. It was common in the early minimalist period to define the strong minimalist thesis by making reference to both sound/sign and meaning, the two external systems that syntax interfaces with. However, it has become increasingly clear in recent years that syntax appears to be designed primarily to interface with meaning. Chomsky puts it thus:

> It may be that there is a basic asymmetry in the contribution to language design of the two interface systems: the primary contribution to the structure of [the] F[aculty of] L[anguage] may be optimization of the C-I [sense] interface. (Chomsky 2008: 136)

The privileged status of meaning over externalization has implications beyond the narrow concerns of syntactic analyses, and is likely to play a prominent role in biolinguistic studies, which focus on the place of language in cognition and its evolution. It has also become clear (though it is rarely made explicit, but see Boeckx in press, Hornstein 2009) that if minimalist research is on the right track, syntax per se is unique in the sense that it is not subject to parametric variation, and furthermore is virtually unaffected by points of variation, which must, by necessity, be relegated to the margin of FL (specifically, the morpho-phonological component of the grammar). Boldly put, minimalist syntax marks the end of parametric syntax (which relies on there being parameters within the statements of the general principles that shape natural language syntax), and leaves no room for an alternative to Borer's (1984) conjecture that parameters are confined to lexical properties. This much should be clear: If minimalist research is taken seriously, there is simply no way for principles of efficient computation to be parametrized. It strikes me as implausible to entertain the possibility that a principle like "Least Effort" could be active in some languages but not in others. In other words, narrow syntax solves interface design specifications optimally in the same way in all languages (contra Baker 2006 and Fukui 2006). I believe that this conclusion is a natural consequence of the claim at the heart of the generative/biolinguistic enterprise that there is only language, Human, and that this organ/faculty emerged very recently in the species, too recently for multiple solutions to design specifications to have been explored.

Thirdly, as Hornstein (2001, 2009) has stressed, minimalism marks the end of grammatical modules. All versions of the P&P model prior to the advent of minimalism took FL to be internally modular, consisting of a variety of largely independent domains, whose combined results yielded surface outputs. By adhering to a radically impoverished notion of FL, minimalists are entertaining the possibility that the core computational system does not contain a variety of internal modules. Rather, simple processes combine with interface conditions to yield surface outputs. This too makes a lot of sense from an evolutionary perspective, given the short window of time we are dealing with in the context of language (see Hornstein in press).

Let me conclude this section by stressing once again that linguistic minimalism is a program, not a theory. A program is a bit like a genome. In isolation it is insufficient to define the shape of an organism, or a theory. It offers certain regulatory principles that constrain the development of an organism/theory. It forces one to pay attention to the role of other factors that come into play during development.

In the context of minimalism, the nature of the external systems with which the core computational system interfaces is all-important, and must be investigated in tandem with syntactic analyses. In this sense, minimalism marks the end of linguistic isolationism, and opens up fresh perspectives for an overall theory of cognition, as I make more explicit in section 19.5.

19.3 A new (radical) chapter in rationalist psychology

Although the plausibility of a minimalist program for linguistic theory was first made explicit in the early 1990s, the suspicion that FL exhibits non-trivial design properties reaches further back. It is, in fact, a fairly natural expectation once we take into account the rationalist roots of the generative enterprise.

I here want to distinguish between two ways of construing the minimalist project. In exploring this minimalist program for linguistic theory, linguists are not just answering the universal Occamist urge to explain with only the lowest number of assumptions (what one might call "pragmatic minimalism", or the weak minimalist thesis), they insist that minimalist analyses really go at the heart of FL as a natural object ("semantic/ontological minimalism", or the strong minimalist thesis.) Principles of well-designed analysis (to be distinguished from principles of good design in nature) have always been part of generative studies, but until Chomsky (1993) no one had thought (explicitly) about elevating principles of good analysis to the ontological or metaphysical level; to move from the nominalist to the realist position. There is a good reason for this. Generative grammarians were concerned primarily with a more immediate problem, the logical problem of language acquisition (what Chomsky 1986*a* called "Plato's Problem"), and the issue of explanatory adequacy. Until some advance could be made on that front, a minimalist program for linguistic theory would have been premature. It was only once the P&P model had stabilized, and been shown how general principles could be extracted and segregated from points of variations (parameters) that it became methodologically sound to try to make sense of these principles in more explanatory terms.

Nevertheless, one can find hints of a minimalist impulse in several early works by Chomsky (for a more extended discussion of these hints than I can afford here, see Freidin and Vergnaud 2001 and Boeckx 2006; see already Otero 1990). As early as 1951, Chomsky wrote (p. 1) that considerations of "simplicity, economy, compactness, etc. are in general not trivial or 'merely esthetic.' It has been recognized of philosophical systems, and it is, I think, no less true of grammatical systems, that the motives behind the demand of economy are in many ways the same as those behind the demand that there be a system at all [Chomsky here refers to Nelson Goodman's work—CB]". This could be paraphrased in a more modern context as "the motives behind the strong minimalist thesis and the search for optimal design are in many ways the same as those behind the demand that there be a (rational(ist)) explanation for properties of FL at all". Short of finding a way of making perfect sense of FL, explanations are bound to be to some degree arbitrary, never going beyond the level of explanatory adequacy. It is perhaps because adaptationist explanations in biology tend to involve a fair amount of contingency, almost by necessity (see Gould 1989, 2002 Monod 1971), that Chomsky has for many years stressed the possibility that "some aspects of a complex human achievement [like language] [may be the result of] principles of neural organization that may be even more deeply grounded in physical law" (Chomsky 1965: 59). For it is in physics that minimalist conjectures have been pursued with much success since Galileo. It is in this domain that one most readily expects good design principles at work. Chomsky's guiding intuition, and the basic contrast between biology and physics, is made clear in this passage from 1982 (Chomsky 1982: 23; but see already Chomsky 1968: 85):

> It does seem very hard to believe that the specific character of organisms can be accounted for purely in terms of random mutation and selectional controls. I would imagine that the biology of 100 years from now is going to deal with evolution of organisms the way it deals with evolution of amino acids, assuming that there is just a fairly small space of physically possible systems that can realize complicated structures.

Citing the work of D'Arcy Thompson, Chomsky points out that "many properties of organisms, like symmetry, for example, do not really have anything to do with a specific selection but just with the ways in which things can exist in the physical world". It seems quite clear to me that Chomsky saw in the success of P&P the very first opportunity to pursue his long-held intuition that the design of FL (and, beyond it, complex biological systems), has a law-like character, determined in large part by very general properties of physical law and mathematical principles.

Already in the 1970s, when he was attempting to unify the locality conditions that Ross (1967) had called "islands", Chomsky sought to explain these "in terms of general and quite reasonable 'computational' properties" (see Chomsky 1977: 89; see also Chomsky 1973: sec. 10), but the degree of specificity of these principles was still quite considerable. And, as Chomsky (1981: 15) notes, considerations of elegance had been subordinated to "the search for more restrictive theories of UG, which is dictated by the very nature of the problem faced in the study of UG" (the logical problem of language acquisition/the search for explanatory adequacy).

Much like one can see hints of linguistic minimalism at work in some of Chomsky's early works and much like one can read Chomsky's most recent writings on the optimal character of language as the systematic exploration of the scattered references to natural laws present already in the 1960s, I think one can see the formulation of the minimalist program as another chapter (albeit a fairly radical one) in Cartesian linguistics, another application of rationalist psychology. When Chomsky wrote *Cartesian Linguistics* in 1966, he was concerned with showing that his own arguments against behaviorism (Chomsky 1959) emphasized certain basic properties of language, such as the creative aspect of language use, or the innate basis of knowledge, that Descartes, Leibniz, and other rationalists of the 17th and 18th century had already identified. Chomsky wanted to stress how much would be lost if these insights were obscured by pseudo-scientific approaches to language such as behaviorism, and how much would be gained by trying to shed further light on the issues that, say, Neo-Platonists brought up in the context of meaning and the nature of concepts (a task that Chomsky has pursued to the present; see Chomsky 2000*b*; McGilvray 2009). Furthermore, he wanted to demonstrate that certain intuitions in the technical works of Port-Royal grammarians matched pretty closely (or could easily be reinterpreted in terms of) what was being done at the time in generative grammar. At the same time, Chomsky was stressing how recent advances in modern mathematics, once applied to language as he had done in Chomsky 1955, 1957, could sharpen some intuitions about the nature of language such as Humboldt's aphorism about language making infinite use of finite means.

Today I think that another aspect of rationalist thought could be said to animate modern (bio-)linguistics, under the impetus of the minimalist program. This aspect pertains to the rationalist conception of life, the sort of conception that was advocated by Geoffrey St Hilaire, Goethe, Owen, and more recently, D'Arcy Thompson, and Turing—those that Kauffman (1993) refers to as the rationalist morphologists.

As Amundson (2005) recounts in his masterful revisionist history of biological thought, the rationalist tradition in biology was obscured not so much by Darwin himself but by all the architects of the modern evolutionary synthesis. Rationalist morphologists had as their main ambition to develop a science of form. They saw development (in both ontogenic and phylogenetic senses) to be governed by laws, revealing a certain unity (of type). They focused on form, and treated function as secondary. They de-emphasized the role of what we would now call adaptation and the power of the environment to shape the organism, and favored internalist explanations according to which development (again, in both its ontogenic and phylogenetic senses) was channeled by physical constraints. Quite correctly, they saw this as the only research strategy to attain a truly explanatory theory of form, a true science of biology. Not surprisingly, they saw it as necessary to resort to idealization and abstraction to reveal underlying commonalities (such as Owen's archetype).

In contrast to all of this, neo-Darwinians, led by Ernst Mayr, focused on function, adaptation, change and the exuberant variety of life. They were empiricists, as is the majority of working (evolutionary) biologists today. But as we will see in the next section, the tide is changing, and laws of form are making an emphatic come-back. There is no doubt that the rationalist morphologists would have been pleased to see the sort of naturalistic, reason-based account of FL advanced in the minimalist program. The attempt to go beyond explanatory is one of the most sustained attempts in linguistics to develop a science of linguistic form, one that abides by the principle of sufficient, or sufficiently good, reason, and views arbitrary aspects of language with the same skepticism with which rationalist morphologists treated contingency.

It may not be too far-fetched to say that just like generative grammar sharpened Humboldt's intuition, minimalist work sharpens the intuition that rationalist morphologists might have had about language. Of course, no one would have been bold enough to regard language as an optimally designed system in the 17th or 18th century, for they lacked the necessary pre-requisite achievement of explanatory adequacy. Though the term Universal Grammar was in use in the 17th and 18th century, the Port-Royal grammarians never went as far as proposing anything remotely like a P&P model, so any attempt to go beyond explanatory adequacy would have been as premature as the minimalist program seemed to be until a decade or so ago. But the minimalist program is imbued by (radical) rationalism, and promises to shed significant light on a key aspect of what makes us human, thereby contributing to the elaboration of Hume's Project of a Science of Man, itself an extension of Descartes' *mathesis universalis*.

Let me conclude this section by pointing out that although I have been at pains to trace back the origin of minimalist thought, I cannot fail to mention that linguistics, even more so than biology, is a very young science. And while it is too easy to forget some of its roots, it is also too easy to forget how truly remarkable it is that already now we can formulate a few minimalist concerns with some precision.

19.4 On the plausibility of approaching FL with Galilean lenses

The approach of the rationalist morphologists touched on above was very "Galilean" in character. Steven Weinberg, who introduced the term into physics, characterized it thus:

... we have all been making abstract mathematical models of the universe to which at least the physicists [read: scientists—CB] give a higher degree of reality than they accord the ordinary world of sensation. (Weinberg 1976)

The Galilean style also characterizes Descartes' work, and is the central aspect of the methodology of generative grammar, as explicitly recognized in Chomsky (1980; see also Hornstein 2005). It is, in other words, another aspect of Cartesian linguistics, or science in general. Galileo notes that

[in studying acceleration] ... we have been guided ... by our insight into the character and properties of nature's other works, in which nature generally employs only the least elaborate, the simplest and easiest of means. For I do not believe that anybody could image that swimming or flying could be accomplished in a simpler or easier way than that which fish and bird actually use by natural instinct. (Galileo 1638 [1974]: 153)

Elsewhere, Galileo states that nature "always complies with the easiest and simplest rules", and that "nature ... does not that by many things, which may be done by few" (1632 [1962]: 99).

The Galilean program is thus guided by the ontological principle that "nature is perfect and simple, and creates nothing in vain" (see, for example, Galileo 1632 [1962]: 397). This outlook is exactly the one taken by minimalist linguists. Indeed, it can be said that the minimalist program is the most thoroughgoing application of the Galilean style in linguistic science, the most radical form of rationalist/Cartesian linguistics yet. The guiding assumption in minimalism is Goedel's basic axiom that *Die Welt ist vernünftig* (the world is full of rationality). The road to Galilean science is to study the simplest system possible, for this is where one is most likely to find intelligibility (rationality). Stephen Jay Gould never tired of emphasizing (see Gould's 1983 delightful essay on explanation in biology "How the zebra gets its stripes"), aesthetic styles (Holton would call them themata; see Holton 1973) profoundly influence the practice of science. Cognitive science is no exception, as Piattelli-Palmarini (1980) already noted in the context of the famous Chomsky-Piaget debate.

The problem for minimalists is that, with its emphasis on the Galilean style of explanation, it has made them look like heretics in Darwin's court. As I mentioned already in the previous section, the modern synthesis in evolutionary biology has not been kind to rationalist thinking and its Galilean style. Whereas it is the style of choice in physics, it has clearly been marginalized in biology. I discuss this clash of scientific styles or cultures at length in Boeckx (2006, ch. 4), so I won't belabor the point here. As Fox-Keller (2002) clearly states, biologists are not sympathetic to idealization, seeing it as a "weakness", a lack of "satisfying explanation" (p. 74), always requiring "more measurement and less theory" (p. 87). Francis Crick (1998: 138) makes essentially the same point when he states that "while Occam's razor is a useful tool in physics, it can be a very dangerous implement in biology". Chomsky himself, already in the early P&P days, was aware of the conflicting outlooks, as he wrote (in a way that highlights once again how latent minimalism was in his earlier writings):

This approach [which Chomsky does not name, but he clearly has the Galilean style in mind—CB], ... is based on a guiding intuition about the structure of grammar that might well be questioned: namely, that the theory of core grammar, at least, is based on fundamental principles that are natural and simple, and that our task is to discover them, clearing away the debris that faces us when we explored the varied phenomena of language and reducing the apparent complexity to a system that goes well beyond empirical generalization and that satisfies intellectual or even esthetic standards.... but it might be that this guiding intuition is mistaken. Biological systems—and the faculty of language is surely one—often exhibit redundancy and other forms of complexity for quite intelligible reasons, relating both to functional utility and evolutionary accident. (Chomsky 1981: 14)

It is for this reason that, although the generative enterprise is firmly grounded in biology, the perspective advocated by minimalists has been deemed "biologically implausible" by many linguists and cognitive scientists alike (see Marcus 2008, Parker 2006, Pinker and Jackendoff 2005, among many others). Jackendoff (1997: 20) nicely sums it up when he says: "it is characteristic of evolution to invent or discover 'gadgets'. (...) The result is not 'perfection'." Jackendoff goes on to say that he would "expect the design of language to involve a lot of Good Tricks (...) that make language more or less good enough. (...) But nonredundant perfection? I doubt it".

I can certainly see how familiarity with popular accounts of evolutionary biology can lead to the claim that linguistic minimalism is biologically implausible, but I have no doubt that the burden of proof will soon shift. Several established figures in biology have started advocating for an enrichment of the standard model in biology (the modern synthesis that emerged some 50 years ago). Gould (2002) made a giant plea for (theoretical) pluralism. For from advocating a wholesale rejection of the Darwinian perspective, Gould stressed the need to recognize non-adaptationist modes of analysis when tackling the problem of form, including the sort of methodology that D'Arcy Thompson's *Growth and Form* represents at its best (see especially Gould 2002: ch. 11).

Although this feeling has not yet reached the popular press, more and more biologists feel that the ultra-adaptationist perspective at the heart of the modern synthesis cannot produce results that qualify as intelligible (read: rational(ist), satisfactory) explanations, and confines biology to a lesser scientific status (making biology "unique"; see Mayr 2004). A growing number of biologists side with Lynch's 2007 opinion that "many (and probably most) aspects of genomic biology that superficially appear to have adaptive roots ... are almost certainly also products of non-adaptive processes". How could it be otherwise, with so few genes (as genomics continues to reveal) for much complexity? Pigliucci (2007*a*) is right to contrast this with the evolutionary psychologists' contention that natural selection should be treated as the default explanation for complex phenotypes; see Dennett 1995, and Pinker 1997, who take Dawkins 1976, 1986 as gospel. I wish they remembered Darwin's claim at the beginning of *The Origin of Species* that "natural selection is . . . not [the] exclusive means of modification" (Darwin 1859 [1964]: 6).

As Carroll (2005) points out, the modern synthesis has not given us a theory of form, the sort of theory that pre-Darwinians were after. But, as the pre-Darwinians recognized, form is so central. As Goodwin and Trainor (1983) write (in a passage that could be lifted from Chomsky's writings), ". . . the historical sequence of forms emerging during evolution is logically secondary to an understanding of the generative principles defining the potential set of forms and their transformations".

Echoing Gould (2002), Pigliucci (2007*b*) is right to say that biology is in need of a new research program, one that stresses the fact that natural selection may not be the only organizing principle available to explain the complexity of biological systems. It is not just all tinkering; there is design too. Pigliucci reviews numerous works that provide empirical evidence for non-trivial expansions of the modern synthesis, with such concepts as modularity, evolvability, robustness, epigenetic inheritance, and phenotypic plasticity as key components.

With minimalist themes in the immediate background, Piattelli-Palmarini (2006) notes that the sort of (adapative) perfection or optimization that neo-Darwinians routinely endorse is just not plausible. There simply hasn't been enough time to optimize organisms gradually. What is needed is "instantaneous" optimization, optimization without search or exploration of alternatives. Empirical results in this domain are coming in, beginning with Cherniak et al.'s (2004) characterization of the neural connectivity of the cortex as the best solution among all conceivable variants. Optima in structures, motions, behaviors, life-styles are now frequently recognized in the pages of *Science* or *Nature*, and none of them seem to be the outcome of long-sought, hard-won, gradualistic adapations. (The latest example of this trend to reach me is a study of the bird's optimal wing stroke [Dial et al. 2008], which vindicates Galileo's claim quoted above that flying could not be achieved in a simpler way than that which the bird uses.) At the same time, the study of biological networks ("systems biology") reveals "special features that give hope that [such] networks are structures that human beings can understand" (Alon 2003: 1,867). Biology is finally yielding to intelligibility. Elsewhere (Alon 2007), Alon writes that biological networks of interactions are simpler than they might have been (or might have been expected to be). Alon clearly states that cells often seem to economize, relying as they do on only a few types of patterns called network motifs that capture the essential dynamics of the system. Alon stresses that approaches that seek to reveal such motifs must rely on abstract representations, focus on the essential aspects, and suppress details—in good Galilean style. Alon is right to conclude his article on "simplicity in biology" by saying that "simplicity in biology must be emphasized" so as to "encourage the point of view that general principles can be discovered". For "without such principles, it is difficult to imagine how we might ever make sense of biology on the level of an entire cell, tissue, or organism" (Alon 2007: 497).

The very same conclusion applies in the domain of linguistic minimalism. Without adhering to the Galilean style, without the strongest possible emphasis on simplicity in language (the strongest minimalist thesis), it is hard to imagine how we might ever make sense of the properties of FL. Chomsky (2004: 124) correctly remarks that "insofar as [minimalism makes progress in capturing properties of FL], the conclusions will be of significance, not only for the study of language itself". If simplicity and efficiency of design are found at the level of the cell and at the level of FL, it is not implausible to expect the same sort of simplicity everywhere in between these two relatively extreme realms of biological structures. Perhaps, then, minimalist pursuits will provide biologists with another model organism in their quest for a science of form.

19.5 The prospects of approaching UG from below: "Applied minimalism"

I have emphasized the metaphysical commitments of linguistic minimalism because it is clear that with the advent of minimalism, linguistics got philosophical, in Whitehead's (1925) sense: "If science is not to degenerate into a medley of *ad hoc* hypotheses, it must become philosophical and must enter upon a thorough criticism of its own foundations." To me, a major part of the excitement of doing linguistics comes from what it reveals about the mind and our species. But I confess that it is now possible to relegate the metaphysical implications of minimalism to the background, and characterize the core methodological principle of minimalist research in a more neutral fashion, by saying (as Chomsky has done in 2007) that minimalism essentially amounts to "approaching UG from below".

This characterization of the minimalist program becomes especially apt when we realize how closely related minimalist research is to the re-emergence of biolinguistic concerns. Chomsky has remarked that

> Throughout the modern history of generative grammar, the problem of determining the character of FL has been approached "from top down": How much must be attributed to UG to account for language acquisition? The M[inimalist] P[rogram] seeks to approach the problem "from bottom up": How little can be attributed to UG while still accounting for the variety of I-languages attained. (Chomsky 2007: 4)

Concretely, approaching UG from below means that the inventory of basic operations at the core of FL must be reduced to a minimum, and much of the richness previously attributed to UG must be re-evaluated: it must be shown to be the result of simple interactions, or else must be attributed to the external mental systems that the core computational system of language interacts with.

Minimalism has forced us to rethink syntax from the ground up (as well as phonology; see Samuels 2009, and semantics; see Pietroski forthcoming, Uriagereka 2008, Hinzen 2007), and find out what is most fundamentally true of, or constitutive of what Hauser et al. (2002) have dubbed the faculty of language in the narrow sense. At the same time, the strongest minimalist thesis requires us to make informed hypotheses about the nature of the external systems that FL serve, which form the faculty of language in the broad sense. As soon as one says that the core computational system of language meets interface demands in an optimal manner, one is forced to adopt an interdisciplinary approach to the study of language. Unsurprisingly, the minimalist program is characterized on the dust jacket of Chomsky (1995) as an "attempt to situate linguistic theory in the broader cognitive sciences".

If indeed much of the specificity of language turns out to be the result of its place in the topography of the mind, it won't do to restrict one's attention to linguistic data to understand language (FL). The systems with which language interacts are bound to be illuminated by minimalist inquiry. Unsurprisingly, questions of meaning, and the relationship between syntactic form and conceptual structures has made an emphatic come-back (see Reinhart 2006, Hinzen 2007, Pietroski 2006, Boeckx 2008*b*, Borer 2005, Uriagereka 2008, Hale and Keyser 2002), as meaning is, in the eyes of many, "the holy grail of the sciences of the mind".

Several authors (see Boeckx 2006, Reuland 2006, Hornstein 2009) have noted that the search for basic principles of organization render FL cognitively and biologically more plausible. Reuland aptly characterizes this state of affairs by saying that the original P&P principles were too good to be false, but much too specific and parochial to be true.

The high degree of specificity of linguistic principles (which I hasten to stress is not specific to the P&P approach pursued by Chomsky and colleagues but is shared by virtually all linguistic frameworks I am aware of: HPSG, LFG, "Simpler Syntax", Relational Grammar, etc.) was necessary to begin to understand the logical problem of language acquisition, but led to a certain feeling that interchanges between linguists and cognitive scientists are "sterile" (see Poeppel and Embick 2005). Poeppel, in particular, stresses that the major obstacle in this respect is the granularity mismatch problem: The degree of abstraction and specificity of linguistic representations and what neuroscientists can today understand is a chasm. As I suggested in Boeckx (2006, ch. 4; see also Hornstein 2009), it is not too implausible to think that the focus on basic, elementary operations and representations in minimalism may help bridge this gap. As Poeppel himself notes, at least some of the operations that minimalists are entertaining (concatenate, merge, copy, linearize, etc.) could conceivably be implemented in neural networks. I don't expect the task to be quick and easy, but I would bet that minimalism has a role to play in turning the mystery of brain implementation into a problem. (Marantz 2005 appears to be equally confident, and deserves credit for resuscitating with the help of his students the much-maligned, but so attractive derivational theory of complexity, which looks plausible again in light of linguistic minimalism.)

Similarly, there is no denying that minimalist constructs allow one to entertain more plausible scenarios concerning the evolution of language than its theoretical predecessors did. Piattelli-Palmarini's (1989) claim that no simple-minded adaptationist account, of the sort put forth by Pinker and Bloom (1990), is likely to be correct still strikes me as exactly right. Piattelli-Palmarini was right to emphasize that many of the properties of language are not amenable to an account that sees communication as the essence of language. But it is fair to say that at the time Piattelli-Palmarini wrote his important essay, the non-adaptationist alternative invoking laws of form didn't look too promising either. How could very general laws of form yield the degree of specificity that the P&P model was made of? Here, too, one faced a granularity mismatch problem. And just as in the case of neuroscience, I think that linguistic minimalism may help us bridge this gap, and make the laws of form conjecture plausible. Here, too, much work remains to be done (and some, like Lewontin 1990, even think that the task is hopelessly beyond our reach), but an article like Hauser et al. (2002; see also Fitch et al. 2005), and the exchanges and studies that it helped foster, I think, demonstrate that progress can be made in this difficult domain.

It is true, as Fitch et al. (2005) remark, that a hypothesis like theirs is strictly speaking logically independent of the minimalist program, but it is hard to see how in practice such a minimal view of the narrow faculty of language as the one they entertain can be seriously put forth without the conceptual and empirical backing up of some minimalist conjecture.

As Wexler (2004) points out in a slightly different context, progress in linguistics—in particular, the advent of a minimalist program—may well turn out to be indispensable in making headway in the two key areas, the logical problems of brain implementation and evolution, that Lenneberg (1967) put at the heart of his project of uncovering the biological foundations of language (biolinguistics). Perhaps minimalists will contribute significantly to making Lenneberg's dream come true.

Let me close this section by pointing out, as Chomsky has done in a very clear fashion in Chomsky (2005), that approaching UG from below relies on the existence of three factors in language design (and indeed, in the design of all organisms, as Lewontin 2000 has stressed): (i) the genetic endowment, (ii) the contribution of the environment, and (iii) principles of growth and laws of form that transcend the limits of "genomic" nativism. It is in its emphasis on these third factor principles that the "bottom-up" approach to UG meets the Galilean style (indeed, the bottom-up approach depends on the plausibility of the latter). It thus turns out that a solution to all the problems of biolinguistics, from Plato's Problem to Darwin's Problem, depends, to a varying degree of radicalness, on rationalist tenets. Indeed, they are unthinkable without what I'd call the Cartesian Program. Without this philosophical foundation, the problems of language acquisition and evolution remain mysteries, as the rationalist philosophers were well aware of. It is no accident that Herder, Humboldt and others built on the foundations of Cartesian linguistics to tackle the problem of the emergence of language in the species; they realized that these foundations offered the only ray of hope in addressing this problem (see Viertel 1966, Schreyer 1985); exactly as minimalists do today.

19.6 Closing remarks

Linguistic minimalism, with its reliance on the Galilean style, and its bottom-up approach that moves away from the standard generative assumption that UG is very rich, may be hard to swallow for many. It is the most radical statement yet that linguistics conducted along rationalist guidelines is not philology by other means. Its focus of attention is not our common sense notion of language, nor is it even grammar or I-language, but only those aspects that fall within the faculty of language in the narrow sense. By stressing the continued reliance on rationalism assumptions, I have tried to indicate that linguistic minimalism is the latest of a series of thought experiments concerning the nature of language that are carried out in an extremely precarious environment (the generative enterprise), fraught with controversy. Although I believe, with Chomsky, that much of the controversy is misplaced, it won't disappear easily (as the resurgence of "new empiricism" and the "rethinking of innateness" makes clear). As Leila Gleitman once said, empiricism is innate. The success of minimalism depends in large part on the success of the generative/rationalist program as a whole. In many ways, linguistic minimalism emerges as the logical conclusion of 50 years of research in generative grammar. By giving its most radical expression to the Generative enterprise, it makes it possible for the very first time to do what Chomsky envisioned in the opening paragraphs of *Syntactic Structures*:

> The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages. (Chomsky 1957: 11)

The minimalist perspective does not, of course, invalidate other approaches to linguistic phenomena; it in fact gives them a certain intellectual coherence and foundation, as all such approaches implicitly or explicitly rely on constructs that must ultimately meet the challenges of acquisition, evolution, and brain implementation. I think that linguistic minimalism, as an addition to the panorama of linguistic analyses, is much needed, and has unique insights to offer into language, its nature, origin, and use.

Further Reading

As in all previous stages of the generative enterprise, Chomsky's writings are required readings to understand the nature of linguistic minimalism. Chomsky's most recent essays have yet to be put together in book form (as were the early essays collected in Chomsky 1995), and as such, they remain to be streamlined, content-wise, but each one of them is indispensable. If I had to single one out, I'd recommend Chomsky 2004 ("Beyond explanatory adequacy"), to be read in conjunction with the less technical Chomsky 2005 ("Three factors in language design"). In addition to Chomsky's works, readers can find overviews of linguistic minimalism ranging from the more philosophical (Uriagereka 1998; Boeckx 2006) to the more technical (Hornstein et al. 2006; Lasnik et al. 2005), as well as a very useful anthology of minimalist studies in Boskovic and Lasnik (2007).

This page intentionally left blank

CHAPTER 20

MORPHOLOGICAL ANALYSIS

GEERT E. BOOIJ

20.1 What is morphological analysis?

Morphology is the subdiscipline of linguistics that deals with the internal structure of words. Consider the following sets of English word pairs:

(1)

Verb	*Noun*
bake	baker
eat	eater
run	runner
write	writer

In these word pairs we observe a systematic form–meaning correspondence: the presence of *-er* in the words in the right column correlates with the meaning component "one who Vs" where V stand for the meaning of the corresponding verb in the left column. The observation of such patterns is the basis for assigning the words in the right column an internal morphological structure $[[x]_V–ər]_N$ where the variable x stands for the phonological form of the base verb. We thus consider these nouns to be complex words. The morphological schema that generalizes over these sets of paradigmatically related words may be formalized as follows:

(2) $[[x]_V–ər]_N$ 'one who Vs'

This schema expresses the systematic form–meaning correspondence found in this set of word pairs. Words are signs with properties at a number of levels of the grammar: they have a phonological form, syntactic properties such as being a noun or a verb, a meaning, and sometimes a particular pragmatic value. Hence, morphology is not a component of the grammar on a par with phonology or syntax. It deals not only with form, unlike what the etymology of the word suggests, but pertains to all levels of the grammar (Jackendoff, 2002). Morphology is the grammar of a natural language at the word level, and calling morphology "the grammar of words" (Booij 2007) is therefore quite appropriate.

The schema in (2) expresses a generalization based on a number of existing verb–nouns pairs of the relevant type. Such schemas also indicate how new complex words can be made. Indeed, the process of creating deverbal *-er*-nouns is quite productive in English. Morphological schemas are word-based since they express generalizations concerning established complex words. In that sense, morphology is word-based. The language user will learn these abstract schemas gradually, after having been exposed to a sufficient number of words that instantiate those schemas. The acquisition of these schemas does not imply that the complex words on which they are based are removed from lexical memory once the schemas have been acquired. Schemas co-exist with the complex words that instantiate these schemas (Bybee 1988*b*, 1995). Hence, the grammar exhibits redundancy, which is no problem given the vastness of human memory. The wrong assumption that the existence of a rule excludes listing outputs of that rule is referred to as the rule-list fallacy (Langacker 1987*b*).

In morphological analysis we also make use of the notion "morpheme", traditionally defined as the minimal meaning-bearing unit of a language. The word *baker*, for instance, might be said to consist of the lexical morpheme *bake* and the bound morpheme *-er*. However, the systematic paradigmatic relationships between words may also be signaled by other means than morpheme concatenation, such as stem alternation, reduplication, stress, and tone patterns. Therefore, the notion "morpheme" is a useful analytic notion for the description of the internal structure of words, but not the starting point of morphological analysis and morphological structure.

The two basic functions of morphological operations are word formation and inflection. Word formation processes create new words, and hence expand the lexicon of a language. Inflection is the grammatical subsystem that deals with the proper form of words in specific syntactic contexts. In Dutch, for instance, the verb *werk* "to work" has five different finite forms, depending on the number and person of the subject of the clause in which this verb occurs:

(3)	werk	present 1 pers.sg
	werk-t	present 2/3 pers.sg
	werk-en	present 1/2/3 pers.pl
	werk-te	past 1/2/3 pers.sg
	werk-te-n	past 1/2/3 pers.pl

We consider these five forms as forms of the same word. The notion "word" in this more abstract sense is usually referred to as "lexeme". Thus, Dutch has a lexeme WERK (lexemes may be indicated by small capitals in order to avoid ambiguity). The stem form of this lexeme is *werk*, and the different inflectional affixes are added to this stem. The word *werker* "worker" is a different lexeme from the word *werk* "to work" (so Dutch has the lexemes WERK and WERKER). The plural form of this noun *werkers* has the following morphological structure:

(4) werk -er -s
work -AGENT -PL
'workers'

This is a simple example of morphological analysis, and presented in a form that follows the conventions of interlinear morphemic glossing (Lehmann 2004). The first line presents the internal constituency of the complex word. The second line provides a morpheme by morpheme glossing, and the third line gives a paraphrase of the meaning of the linguistic unit.

The set of verbal forms in (3) illustrates the well-known problem that there is no one-to-one mapping between morphemes and units of (grammatical) meaning, also referred to as "cumulative exponence". For instance, the *-t* of *werkt* expresses values for the following grammatical categories:

(5) Person: 2 or 3
Number: singular
Tense: present

For this reason, a morpheme like *-t* is traditionally called a "portmanteau morpheme". The Dutch sentence *Jan werkt* will receive the following glossing:

(6) Jan werk-t
John.3SG work-3SG.PRES
'John is working'

Grammatical features that are expressed by the same morpheme are separated by a dot instead of a hyphen. The combination of feature values for person and number is usually given without an internal dot.

Two other notions are important for morphological analysis, the notions "root" and "stem". The stem of a word is the form minus its inflectional markers. The root of a word is the stem minus its word formation morphemes. Hence, in the English word *workers* the stem is *worker*, and the root is *work*. Another example is that in Polish the root noun *kos* "scythe" can be turned into a verbal stem with the meaning "to mow" by adding the verbalizing suffix *-i*. This verbal stem can then be used for deriving verbal forms such as the present participle *košonc* "mowing", the phonetic form of *kos-i-onc*.

20.2 Word formation

Natural languages make use of a number of formal means for the formation of complex lexemes: compounding, affixation, reduplication, conversion, stem alternation (also referred to as internal modification), stress, and tone.

In compounding two or more lexemes are combined into a new one. Cross-linguistically, compounding is one the most common means for word formation, in particular compounding in which one of the constituents is the head (so-called endocentric compounding). The English word *football* is a compound, consisting of the two lexemes *foot* and *ball*, of which the second functions as the head: a football is a particular kind of ball, not a kind of foot. There are also languages with left-headed compounds, such as Maori. The Maori compound *roro-hiko* “lit. brain electricity, computer” denotes a particular kind of brain, namely a computer, not a particular form of electricity. In exocentric compounds such as *pickpocket* there is no constituent that functions as the head: a *pickpocket* is neither a pocket nor a pick. An example of an excocentric compound from Mandarin Chinese, a language with lots of exocentric compounds, is the compound *tian fang* consisting of the verb *tian* “to fill” and the noun *fang* “room” with the meaning “second wife (to a widower)”. Besides subordinating compounds, with one of the constituents functioning as the head, there are also coordinating compounds, such as Sanskrit *maa-pio* “mother and father”, and English *singer-actor*.

The second widespread process used in word formation is affixation, whereby an affix is prefixed, suffixed, infixed, or circumfixed to some input form. Each of the four options is illustrated in (7):

(7)	prefixation:	*un-happy* from *happy* (English)
	suffixation	*happi-ness* from *happy* (English)
	infixation	*s-m-ka:t* ‘to roughen’ from *ska:t* ‘rough’ (Khmu, a language spoken in Laos)
	circumfixation	*ge-been-te* ‘bones’ from *been* ‘bone’ (Dutch)

Compounding and affixation are referred to as concatenative morphology since their mode of operation is that of concatenating roots, stems, and affixes. A special form of concatenative morphology is reduplication. This is the process in which a stem or part thereof is copied and prefixed or suffixed to that stem, as in Javanese *baita-baita* ‘various ships (full reduplication) and *tə-tamu* “to visit” from *tamu* “guest” (partial reduplication with copying of the initial consonant and insertion of a default vowel).

In non-concatenative morphology, other formal means are involved in the creation of new morphological forms. In the case of internal modification a stem with a different form is created, for instance, by replacing a vowel pattern or a consonant pattern (or both) with another one. Vowel alternations are characteristic of a number of Indo-European languages; in Semitic languages verbal roots may appear in a number of different "binyanim", templates with specific patterns of consonants and vowels, sometimes in combination with a prefix:

(8) *stem alternation* (Dutch, Indo-European)
sl*ui*t 'to close' sl*o*t 'lock'
b*i*nd 'to bind' b*a*nd 'bond'
binyan system (Modern Hebrew, Semitic)
katav (pattern CaCaC) 'wrote'
ni-ktav (pattern ni-CCcaC) 'was written'
kitev (pattern CiCeC) 'inscribed' (intensive meaning)

Other non-concatenative means for marking morphological operations are stress (as in the English word pair *to revíew* (verb) vs. *réview* (noun) and the use of tone to mark specific morphological categories. In some cases the tonal marking can be analyzed as a case of concatenative morphology. An example of the latter from the African language Noni is the following set of pairs of singular and plural nouns (Hyman and Leben 2000, p. 590):

(9)

singular: LH tone pattern	*plural: H tone pattern*	
bwě	bwé	'dog'
jǐn	jín	'maggot'

Since the roots involved may be assumed to have a lexical High tone, Hyman and Leben qualify the Low tone that is part of the singular tone pattern as a tonal prefix that marks the singular. Hence, affixes can also consist of suprasegmental units.

New lexemes may also be created without overt formal marking, which is referred to as conversion. A well-known case is the conversion of nouns into verbs, a very productive process in languages like English and Dutch; the following examples are from Dutch (the nouns are recent English loans except *contact*):

(10)

noun	*verb*
contact 'id.'	contact 'to make contact with'
computer	computer 'to make use of a computer'
skype	skype 'to communicate by means of Skype'
sms	sms 'to send an SMS message'

These recently coined examples show how productive this way of creating lexemes is in Dutch. The morphological structure of such verbs can be represented as $[[x]_N]_V$, and the corresponding meaning as "to perform an act in which N is involved". In this way we avoid the assumption of arbitrary zero-morphemes that are sometimes used in morphological analyses to account for conversion.

The common denominator for all word formation processes except compounding is derivation. Hence, the notion "word formation" comprises both compounding and derivation.

20.3 The lexicon

The term "lexicon" refers to the component of the grammar that minimally contains a specification of the lexical units of a particular language. The set of lexical units is larger than the set of words. Idiomatic expressions such as *to kick the bucket* "to die" that are phrasal in nature need to be listed as well. This also applies to the many noun phrases that are used as classifying terms such as *blue cheese* and *yellow pages* that form established ways of denoting certain entities. The distinction between the notions "word" and "lexical unit" is important for a proper understanding of the relationship between morphology and syntax, as we will see below.

The set of words that are to be registered in the lexicon is the set of established words, that is, the set of words that is used by more than one native speaker and on more than one occasion. Thus, the lexicon is part of the language norm since it specifies the lexical conventions of a language. This norm can be changed by adding new words to the lexicon.

New complex words, once established, will be added to the lexicon. From a diachronic perspective there are other means of extending the set of complex words as well, as will be discussed below, in the section on diachrony.

In the lexicon, established complex words coexist with the schemas according to which they are formed. The schemas express generalizations about sets of established words, and indicate how new words can be formed. The relation between the schemas and their instantiations can be conceived of as a hierarchical lexicon in which a schema forms a node that dominates its instantiations. All properties shared by a set of words are specified in the schema, and the individual words inherit these properties from the node that dominates them.

20.4 Inflection

So far, the focus of this chapter has been on lexeme formation. In this section, we will consider the analytical challenges posed by the other domain of morphology, the study of inflectional systems. In a language with inflection, lexemes may have more than one inflectional form. The set of inflectional forms of a lexeme is traditionally represented in the form of a paradigm in which each cell contains a form that expresses a particular array of grammatical features. Inflection has two basic functions. The first is that of creating forms of lexemes with a formal marking for certain grammatical categories such as number for nouns, and tense and aspect for verbs. This kind of inflection is referred to as inherent inflection (Booij 1993, 1996) because the choice of the inflectional form is not governed by syntactic context but by semantic considerations; it reflects the choice of the language user as to what information he or she wants to convey. The other kind of inflection is contextual inflection, inflection that is determined by the syntactic context in which a lexeme occurs. In a language with a case system, for instance, nouns must have a particular case form because they are the head of an NP in a certain syntactic position that requires case marking. In many languages, finite forms of verbs have to agree with person and number properties of the subject NP. Adjectives may have to agree with respect to certain properties (such as gender, case, and (in)definiteness) with the nominal head that they modify. Verbs and prepositions may require specific case markings on their NP-arguments, as is the case for German. The following sentence illustrates both inherent and contextual inflection for Latin. The cases of pure contextual inflection are in bold print:

(11) Te semper, ut omn-**ibus** pate-t,
You.ACC.2SG always, as all-DAT.PL is.clear-PRES.3SG
immoderat-**o** amor-e
unfettered-MASC.SG.ABL love-MASC.SG.ABL
complex-**a** sum
embrace.PART.PERF-FEM.SG.NOM be.PRES.1SG
'I have always embraced you with unfettered love, as everyone knows.'

(DAT = dative, ABL = ablative). This sentence is from a famous medieval love letter by Heloïse to Abelard (Janson 2004, p. 139). The verb *patere* "to be clear" requires its non-subject argument to be marked by the dative case. The ablative marking on *amor* "love" is a case of inherent inflection ("semantic case"), chosen to express a circumstantial meaning. The corresponding marking on the adjective *immoderato*, on the other hand, is a case of contextual inflection, required by the rule of agreement between modifying adjective and nominal head. The word sequence *complexa sum* is the perfective form of a so-called deponent verb, a verb with active meaning but a passive form. In the perfect tense, such deponent verbs have a periphrastic form, consisting of two words, a perfect/past participle and a form of the verb *esse* "to be". The participle has inflection for feminine gender since the writer of this sentence, Heloïse, is a female. The subject itself, however, is not expressed by a separate noun. We may still consider this contextual inflection if we take the notion "context" to include the pragmatic context. The ending -*a* expresses both feminine gender and nominative case. An overt subject for the finite form *patet* is missing as well. In fact, the form of the verb enables us to reconstruct the subject as a 3rd person singular entity. The verb thus agrees in person and number with an abstract, non-overt subject. Hence, the suffix -*t* of *patet* expresses both inherent inflection (tense) and contextual inflection.

The word *sum* illustrates another morphological phenomenon, that of suppletion. This is the situation of different lexical roots playing a role in filling the cells of the paradigm of a lexeme. Two additional formal complications in inflection are the role of inflectional classes (declinations for nouns and conjugations for verbs), and stem selection. In example (11), the word *amor* "love" belongs to the "third declination" of nouns which implies that the ablative singular case is expressed by the ending *-e*, whereas the participle *immoderat-us* is inflected according to the default declination class for adjectives, hence the corresponding ending *-o*. In verbal paradigms the stem to which the inflectional endings are attached may have more than one form. The verb *patet* is a form of the verb *patere* "to be clear" that belongs to the second conjugation, the verbs with the "thematic vowel" *e* in between the root and the inflectional endings. The stem form of the participle *complexa* is /pleks/, whereas the form /plekt/ is used in the finite forms present and past (as in *com-plect-or* "I embrace").

This small piece of morphological analysis, which includes the use of notions like suppletion, deponent verb, stem allomorphy, and inflectional class, shows how complex the relation between form and meaning can be in inflectional systems (Aronoff 1994).

The contextual inflectional marking on the Latin adjective *immoderato* is a case of dependent marking since the adjective is dependent on the noun, the head of the noun phrase *immoderato amore*. There are also languages that mark the head rather than the dependent (Nichols 1986). For instance, Hungarian exhibits head marking, as in:

(12) az ember ház-a
the man house-3SG
'the man's house'

where the noun *haz* "house" is the head.

The functional distinction between inherent and contextual inflection made above can be used for predicting the order in which the relevant inflectional elements occur in complex words: contextual inflection tends to be peripheral to inherent inflection. For instance, the ablative singular form of the Finnish word for "cat" is *kisso-i-lta*, with the ablative suffix *-lta* ordered after the plural suffix *-i*. Inflection in its turn is peripheral to word formation. When inflectional systems erode, it is usually the contextual inflection that gets lost first. For instance, most Romance languages have lost their case system while preserving the morphological expression of number on nouns, a case of inherent inflection.

The existence of a rich inflectional system implies that a lexeme may have quite a number of forms. Unlike what is the case for lexeme formation, it is therefore not very realistic to assume that all inflected forms of the established lexemes of a language are stored in the lexical memory of speakers, certainly not for languages with rich inflection. Storage of inflected forms is most probable for irregular forms and for regular forms with a certain frequency of occurrence (Booij 1999*a*).

20.5 Interfaces

20.5.1 The interface with phonology

As we saw above, the basic levels of morphological analysis are the phonological level, the level of morphosyntactic structure, and the semantic level. The interface between these three types of information is subject to certain general principles.

As to the interface between phonology and morphology, the morphological structure of a complex word co-determines the phonological form of the complex word, in particular its prosodic structure. In the default case, a word in the morphological sense corresponds to a word in the phonological sense, referred to as the phonological word or the prosodic word. For instance, the English word *baker* is one prosodic word. The prosodic word is the domain of syllabification (that is, the division of the phonological string of a word into syllables). The syllabification of the word *baker* is as follows:

(13) (be:.kər)$_\omega$

(the dot indicates a syllable boundary, the ω stands for "prosodic word"). This shows that morphological structure and prosodic structure are not necessarily isomorphic: there is no prosodic boundary that corresponds to the word-internal morphological boundary of this word. Since the suffix *-er* forms one prosodic word with the stem, we call it a cohering affix. Affixes may also form a prosodic word of their own, however, and are then qualified as non-cohering. Prefixes in Germanic languages are often non-cohering, and a number of suffixes as well. For instances, in careful speech the English word *un-able* has a syllable boundary before the first vowel of *able* even though normally a consonant belongs to the same syllable as the next vowel. Hence, *un-* is a non-cohering prefix (Booij and Rubach 1984), and forms a prosodic word of its own. Therefore, *unable* is a prosodic compound. Suffixes may also be non-cohering. An example is the Dutch suffix *-achtig* "-like", as in *rood-achtig* "red-like, reddish". The prosodic structure of this word is (ro:d)$_\omega$ (ɑx.təɣ)$_\omega$. Consequently, the final /d/ of the first constituent *rood* with the underlying phonological form /ro:d/ is devoiced since Dutch obstruents are voiceless at the end of a syllable, and the phonetic form of this word is [ro:t.ɑx.təx]. Thus, the phonetic form of this adjective contrasts with that of the synonymous adjective *rod-ig* [ro:dəx] with the cohering suffix *-ig* /əɣ/ in which the voicedness of the underlying /d/ is preserved because it does not occur in syllable-final position. In many languages the lexeme constituents of compounds form separate prosodic words, with the effect that the compound-internal morphological boundary coincides with a syllable boundary. Thus we get audible minimal pairs of the following type (examples from Dutch):

(14) loods-pet 'pilot cap' [lo:ts.pɛt] lood-spet 'lead drop' [lo:t.spɛt]
balk-anker 'beam brace' [bɑlk.ɑŋ.kər] bal-kanker 'testicle cancer' [bɑl.kɑŋ.kər]

A second example of the influence of morphological structure on the phonetic realization of complex words is that of word stress, one of the most intensively studied aspects of phonology as far as English is concerned. Some English suffixes are stress-neutral, that is, they do not change the stress pattern of their stem, whereas others are stress-shifting, and shift the main stress of the stem rightward. The English suffix *-er*, for instance, is stress-neutral, unlike the suffix *-ee*, as can be seen in a pair like *emplóy-er* vs. *employ-ée* (cf. the classical study of Chomsky and Halle (1968), and for a more recent discussion Hammond (1999)). The interaction between phonology and morphology is the main focus of the theory of Lexical Phonology (cf. Booij (2000) for a survey), and also an important domain of research in Optimality Theory (Kager 1999).

20.5.2 The interface with semantics

The basic principle that governs the interface between the formal structure of complex words and their meaning is the principle of compositionality: the meaning of a complex word is a compositional function of the meaning of its constituents and the morphological structure. For instance, the meaning of *football* is a compositional function of the meanings of *foot* and *ball*, and the meaning contribution of the compound structure that can be circumscribed as follows for English compounds, which are right-headed:

(15) $[X\ Y]_Y$ 'Y that has some relation R to X'

Hence, a football is a ball which has something to do with feet. The exact relation between *foot* and *ball* does not belong to the domain of linguistic knowledge but is part of our knowledge of the world (of games). We know that the foot is the device used for hitting the ball. This makes clear that the specific semantic interpretation of individual words is underdetermined by the linguistic system as such.

A second form of interface in which semantics is involved pertains to argument structure. The creation of complex verbs may have predictable consequences for the syntactic valency of these verbs. For instance, when we create causative verbs from adjectives, with the meaning "to cause something to become A", the event expressed by the causative verb presupposes an Agent and a Patient. Hence, such causative verbs will have at least two arguments, and hence be transitive. Thus, the semantics of a class of complex words may have predictable consequences for their syntactic valency. Such rules that predict the relationship between argument structure and syntactic valency are referred to as "linking rules".

The principle of compositionality takes a syntagmatic perspective on the semantics of complex words. There are clear cases of word formation, however, where we need a paradigmatic perspective in order to account for the meaning of a complex word. The meaning of the Dutch compound *huisman* "househusband", for instance, can only be understood when seen as part of the following equation:

(16) vrouw 'woman': man 'man' = huisvrouw 'housewife': huisman 'househusband'

A *huisvrouw* is a woman without a paid job who stays at home to take care of the household, a *huisman* is the male counterpart of such a woman.

The idea that paradigmatic relationships play a role in the semantics of complex words can also be seen in the interpretation of Dutch words ending in the noun *boer* "farmer". This word functions as the head of a compound like *groenteboer* "green-grocer", reflecting a time when the farmer was both the grower and the retailer of vegetables. In present-day Dutch, the constituent *boer* has developed the more general meaning retailer, witness compounds like *sigarenboer* "cigar seller, tobacconist" and *tijdschriftenboer* "magazines salesman". Thus, Dutch developed a particular subpattern of NN compounds of the following type:

(17) $[[\text{x}]_N\ [\text{boer}]_N]_N$ 'seller of x'

which may be qualified as a "constructional idiom" (Jackendoff 2002) at the word level: a construction of which one position is lexically filled and another one still a variable.

The same observation can be made for the left constituents of compounds. For instance, in most Dutch compounds that begin with the word *hoofd* "head", the meaning of that constituent is "main", as in *hoofd-ingang* "main entrance" and *hoofd-bezwaar* "main objection". In Maale, an Omotic language spoken in southern Ethiopia, the noun *nayi* "child" has developed into a word with the general meaning of agent, as in *bayi nayi* lit. "cattle child, one who brings cattle to grazing area". This reflects the fact that cattle herding is typically a children's task in that society (Amha 2001, p. 78). Thus, such compound constituents may develop into affix-like elements (referred to as affixoids).

This implies again that we conceive of the lexicon as a hierarchy of levels: at the bottom the individual coined complex words, at the top the abstract schemas according to which these words have been formed, and intermediate generalizations like (17). In the case of NN compounds, we thus get (at least) three levels):

(18) $[[\text{x}]_N\ [\text{y}]_N]_N$ 'y with some relation R to x'

|

$[[\text{x}]_N\ [\text{boer}]_N]_N$ 'seller of N'

|

$[[\text{sigaren}]_N\ [\text{boer}]_N]_N$ 'seller of cigars'

At each level a construct(ion) instantiates the constructional schema by which it is dominated.

20.5.3 The relation between morphology and syntax

Our approach so far can be qualified as "lexicalism". This term denotes the set of theories in which morphology is separated from syntax in the sense that the structure of complex words is not dealt with by the syntax but by lexical rules that express generalizations about established and potential complex words. This does not mean that syntax and morphology do not interact, but that syntactic rules cannot manipulate parts of words. This principle is referred to as the principle of Lexical Integrity:

(19) *Principle of Lexical Integrity*
'The syntax neither manipulates nor has access to the internal structure of words' (Anderson 1992, p. 84)

As argued in Booij (2009), this principle is too strong, since one should not exclude the possibility that syntactic rules may have access to the internal morphological structure even though manipulation must be excluded. An example that shows that access of syntactic or semantic rules to word-internal structure cannot be completely ruled out comes from Georgian. In Georgian we find expressions such as the following (Harris 2006):

(20)	sam	tit-moč'r-il-i	(k'aci)
	three.OBL	finger-cut.off-PTCP-NOM	man.NOM
	'(a man) with three fingers cut off'		

The first word *sam* has to appear in the oblique form because it modifies the word *tit* "finger" which is part of the second word. That is, both for the purpose of case assignment (to the independent word *sam* only) and semantic interpretation, *sam* and *tit* form a unit. As Harris argues, the word *sam* cannot be considered a part of the next word even though its form is indeterminate since it could also be a stem form, because recursive modification is not allowed within Georgian compounds. Hence, it should be interpreted as the oblique form of an independent word. This case assignment thus requires access to the internal morphological structure of the second word in (20). The construction in (20) may be compared to that in (21) where the first word bears nominative case, and you get a different interpretation:

(21)	sam-i	tit-moč'r-il-i
	three-NOM	finger-cut.off-PTCP-NOM
	'three (men, people, statues) with fingers cut off'	

In (21), the word form *sami* agrees in case marking with the second word as a whole, and hence it is a modifier of the whole word. Note, however, that the word *tit*, being part of a compound, does not receive case marking itself.

The need to access the internal structure of complex words is also shown by scope phenomena: in some cases a modifier may have scope over a sub-constituent of a complex word. This is illustrated in (22) with Dutch phrases in which the pre-nominal adjective modifies only the first noun constituent of the NN compounds. I use Dutch examples here even though the English glosses have the same properties. For Dutch we can be certain that these linguistic units are phrases because the adjectives are inflected, witness the inflectional ending -*e*:

(22) $[\mathrm{A}\ [\mathrm{NN}]_{\mathrm{N}}]_{\mathrm{NP}}$

visuel-e informatie-verwerking
visual-NONNEUTER.SG.INDEF information-processing
'visual information processing'

intellectuel-e eigendoms-rechten
intellectual-NEUTER.PL.INDEF property-rights
'intellectual property rights'

The principle of Lexical Integrity can be used to determine the status of lexical units such as preverb + verb combinations that are found in several Indo-European languages, and also in Uralic languages like Hungarian (Kiefer and Honti 2003). The latter language has lexical units such as *tévét nez* "be engaged in television watching". The two parts of this lexical unit can be split in certain syntactic contexts, for instance by the negative word *nem* "not". The splittability of these units is evidence for their not being words. This is confirmed by the fact that the noun constituent *tévét* in this example is marked with accusative case by the suffix -*t*. This assignment of structural case to *tévé* shows that it must be an independent word. Given the principle of Lexical Integrity, one does not expect structural case assignment to a sub-constituent of a word. Thus, we can make a principled distinction between morphological and syntactic constructs.

This does not imply that syntactic constructs cannot form part of words. In the following Dutch examples, AN phrases are used in the non-head position of complex words. These complex words as such are morphological constructs, but one of their constituents is formed in accordance with the rules of syntax:

(23) $[[[\text{oude}]_{\mathrm{A}}\ [\text{mannen}]_{\mathrm{N}}]_{\mathrm{NP}}\ [\text{huis}]_{\mathrm{N}}]_{\mathrm{N}}$ 'old men's home'
$[[[\text{vierde}]_{\mathrm{A}}\ [\text{klas}]_{\mathrm{N}}]_{\mathrm{NP}}\ \text{-er}]_{\mathrm{N}}$ 'fourth class pupil'

These cases of word formation do not invalidate lexicalism since the distinction between syntactic and morphological constructs is maintained.

20.6 Morphological classification

Languages may be classified according to the role and nature of morphology in each language (Comrie 1981; Haspelmath 2009c). A first dimension of classification is the index of synthesis: languages that do not make use of morphology are called analytic or isolating, languages with a lot of morphology are called synthetic. Hence, languages may be ranked on an index of synthesis. Traditionally, Chinese is referred to as an isolating language because it has no, or almost no inflection. Yet, there is no doubt that word formation, in particular compounding, is very productive in this language (Packard 2000). Hence, Chinese is not analytic in an absolute sense.

The second index on which languages can be ranked is that of polysynthesis: some languages allow the incorporation of stems, leading to relatively complex words, as illustrated by the following one-word-sentence of Central Alaskan Yup'ik (Mithun 2000) p. 923):

(24) Tuntutuq=gguq
tuntu-te-u-q=gguq
moose-catch-INDIC.INTRANSITIVE=HEARSAY
'He got a moose.'

The third dimension of classification is the index of fusion. In fusional languages, one morpheme may express more than one grammatical feature. Above, we saw that Latin is such a language. Such languages can be contrasted with agglutinating languages in which each bound morpheme corresponds with one grammatical feature. Turkish is the textbook example of an agglutinating language. For instance, case and number in Turkish are expressed by different suffixes, unlike what is the case for Latin:

(25) çocuk-lar-ın
child-PL-GEN
'of the children'

These three indices of morphological complexity are useful in given a global characterization of the morphology of a language. One should be aware, however, that languages are not homogeneous with respect to these indices (Haspelmath 2009c). For instance, many Indo-European languages are fusional in their inflectional system but agglutinating in their derivational morphology. Chinese also illustrates this point since, as mentioned above, it is synthetic as far as word formation is concerned but analytic as far as inflection is concerned.

20.7 Affix ordering

In languages with a reasonably rich morphology affix ordering is an important topic for morphological analysis. The basic question is how we can account for the ordering in which the different types of morphemes have to appear in multiply complex words.

A well-known general principle (Greenberg 1963) is that inflectional morphemes are peripheral to derivational morphemes. Within the domain of inflection, contextual inflection appears to be peripheral to inherent inflection (Booij 1993, 1996).

The ordering of inflectional affixes has also been investigated in detail by Bybee (Bybee 1985*a*). Bybee proposed a semantic Relevance Hierarchy for the order of affixes: the more semantically relevant an affix is for the stem, the closer it is to the stem. Hence, derivational morphemes, which have obviously a profound effect on the meaning of the stem are closer to the stem than inflectional ones.

As to the inflectional affixes on nouns, it is predicted that case markers and (in)definiteness markers on nouns will be peripheral to gender markers since they do not have a semantic effect on the stem of the noun. Instead, they relate the noun to its syntactic context.

In the case of verbs, the following hierarchy can be observed for languages with markers for voice, aspect, tense, and agreement:

(26) Voice > Aspect > Tense > Agreement

Voice markers such as passive have a strong semantic effect, and affect the argument structure of the verb. At the other end of the hierarchy, tense has the deictic role of relating the event expressed by the verb to the moment of speaking. Agreement, a case of contextual inflection, also has an external role in that it relates the verb to its syntactic context.

Another example is the ordering of suffixes in the following Maale verb (Amha 2001, p. 114):

(27) gap-is-é-ne
finish-CAUSATIVE-PERFECT-DECLARATIVE
'finished'

The causative suffix affects the argument structure and hence the syntactic valency of the root *gap*. Hence, it is closest to the root. The declarative suffix on the other hand expresses a property of the whole sentence since it indicates that the sentence is declarative. That is, it does not modify the semantic content of the verbal root as such and is therefore the most peripheral suffix.

For some languages with complicated sequences of morphemes in words one finds descriptions with templates. A template specifies a number of slots for specific morphemes. This kind of affix ordering is referred to as "position class morphology" (cf. Inkelas (1993) for a discussion of such template morphology in Nimboran, a Papuan language of New Guinea). For Athapaskan (Amerindian) there is a detailed study that relates the order of affixes to their semantic scope properties (Rice 2000).

The order of affixes in a complex word may also reflect the different historical strata of the vocabulary of a language. The Dutch vocabulary, for instance, has a non-native (Romance) stratum besides a native (Germanic) stratum. The basic generalization is that non-native suffixes can attach to non-native (simplex or complex) stems only, whereas native suffixes can attach to both non-native and native stems (Booij 2002*b*). Hence, the predicted order is that native suffixes will be peripheral to non-native ones, as illustrated in (28):

(28) real-iser-ing 'realization'
controvers-ieel-heid 'controversialness'

The suffixes -*iseer* and -*ieel* are non-native suffixes borrowed from French, whereas the suffixes -*ing* and -*heid* are native suffixes of Germanic origin. The non-native suffixes first attach to the (non-native) roots. They cannot be added after a native suffix, since the attachment of a native suffix makes the stem native.

As to English, there is a long debate on how to deal with the constraints on suffix ordering (Hay and Plag 2004; Plag 1996). This is related to how complex words are processed, discussed in section 20.10 below. The idea is that affixes that are easily and often recognized as parts of complex words tend to be peripheral to affixes that form part of complex words whose internal structure is not so easily parsed. For instance, the English suffix -*less* is easily parsed out, and attaches freely to all kinds of complex words, whereas -*ity* that occurs in less parsable words has a more restricted distribution. Hence, a word like **home-less-ity* is odd though it is semantically wellformed (Hay 2002).

Prosodic properties may also play a role in stacking up affixes. Dutch suffixes that are non-cohering and thus form prosodic words of their own, can easily be attached to already suffixed words—sometimes even to plural forms of nouns—unlike cohering suffixes. For instance, the productive cohering suffix -*ig* /əɣ/ "-ish" cannot attach to adjectives that have a participial form whereas the non-cohering suffix -*heid* "-ness" freely attaches to such participial adjectives. Hence, the contrast in wellformedness between **woed-end-ig* "slightly furious" and *woed-end-heid* "furiousness". That is, words can be made longer if the suffix starts a new prosodic word (Booij 2002*a*).

20.8 Diachrony

The use of productive word formation patterns is not the only source of complex words in the lexicon. Complex words may also arise through univerbation, the process in which phrases become words. Many nominal compounds in Germanic languages have a phrasal origin. For instance, the Dutch compound *koningskroon* "king's crown" originated as a phrase in which the noun *koning* was marked as the possessor through the genitive case ending *-s*. The case ending that was thus trapped inside a word was then reinterpreted as a semantically empty linking element or stem extension. The system of linking elements became subsequently part of the compounding system of Dutch.

Word formation processes have the function of expanding the sets of words of lexical categories (nouns, verbs, adjectives, and adverbs). Yet, we also find complex words of non-lexical categories such as prepositions. This is due to the process of grammaticalization, defined as follows by Hopper and Traugott (2003, p. xv): "We define grammaticalization as the process whereby lexical items come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions." The English preposition *during*, for instance, has the shape of a present participle of the verbal root *dure* that we also find in *en-dure* and *dur-ation*. The participle could be reinterpreted as a preposition in absolute participial constructions like *during the war* "while the war lasted". Thus, the class of English prepositions (prepositions are grammatical morphemes) was expanded with a complex word *during*. In the word *notwithstanding* we see a combination of univerbation and grammaticalization.

Grammaticalization can lead to the rise of word formation processes since lexemes can become bound morphemes, prefixes or suffixes, which belong to the class of grammatical morphemes. In French, for instance, the French preposition *sur* has a prefixal counterpart *sur-* with the meaning "over", as in *surexposition* "overexposure" (Amiot 2005). The English suffix *-wise* as used in *money-wise* "as far as money is concerned" originates from the noun *wise* "manner". Thus, univerbation and grammaticalization are mechanisms of language change that lead to the expansion of non-lexical categories, and to the rise of new derivational processes (Booij 2005*a*; Heine and König 2005).

Complex words can also be subject to the process of lexicalization, and thus lose their morphological transparency. The Dutch word *aardappel* "potato", for instance, is historically a compound, consisting of the stems *aard* "earth" and *appel* "apple". Yet, it is no longer perceived as a kind of apple, and it is syllabified as a simplex word, without the word-internal morphological boundary coinciding with a syllable boundary: *aar.dap.pel*, not *aard.ap.pel*.

Language contact is another source of word formation processes. Dutch, English, and German have borrowed many complex words from French in the course of time, for instance deadjectival nouns ending in *-ité*, with some phonological adaptation (Dutch *-iteit*, English *-ity*, German *-ität*). Speakers of these languages could abstract a word formation schema on the basis of a number of such loans and the corresponding adjectives, and use these productively. Thus, Dutch now has a number of nouns in *-iteit* for which there is no French counterpart since they have been coined in Dutch on the basis of a Dutch adjective, such as *stommiteit* "stupidity" derived from *stom* "stupid". Similarly, the English deverbal suffix *-ee* derives from the French passive participle ending *-é*, but has now gone its own way and combines with English verbs, as in *standee* derived from the verb *stand*.

Word formation processes can also be lost. In the course of time a word formation process can lose its productivity, with the effect that no more words of that type are coined. This means that we might still find a number of instantiations of the relevant word formation schema, but no extension of that class. The Dutch suffix -*lijk* for instance, which is similar in function to the English suffix -*able* in coining deverbal adjectives, lost its productivity, and the role of coining new deverbal adjectives is taken over by the productive suffix -*baar* "-able".

As mentioned above, inflectional systems may be subject to strong erosion. Most Romance and Germanic languages have lost their case system for nouns and agreeing adjectives. Verbal inflection may also change considerably. Afrikaans, a daughter language of Dutch spoken in South Africa, lost most of its verbal morphology, with the effect that there is only one verb form left. Past tense in this language is expressed periphrastically, as illustrated in (29):

(29) *Ek het geloop*
I have walked
'I walked.'

This kind of inflectional erosion is due to the effect of contact between speakers of Dutch and speakers of African languages and Malay in South Africa. The simplification of the inflection of nouns in English (loss of case and gender marking) may also be due to the effect of language contact between Anglo-Saxon speakers and Viking invaders.

Language change may lead to the rise of constructions with specific morphology. Consider the following English phrases, and their labeling (Rosenbach 2006, 2007):

(30) Determiner genitive: John's book, the young girl's eyes
Descriptive (classifying) genitive: a men's suit, the lawyer's fees

The morpheme -*s* in these constructions derives historically from a genitive case ending. After the loss of the case system, this use of -*s* for the marking of specifiers in certain types of noun phrases persisted. Hence, this use of the -*s* is a case of construction-dependent morphology (Booij 2005*b*). Another example is the use of the old dative suffix -*en* in Dutch collective constructions such as

(31) met zijn vier-en na en-en
with his four-*en* after one-*en*
'the four of us' 'after one o'clock'

This use of the suffix -*en* does not follow from synchronic case marking but is the residue of case marking in an earlier stage of Dutch. Morphological elements can be "trapped" in a construction, and thus become dependent on that construction.

As mentioned above, individual complex words may lose their internal morphological structure in the course of time, a form of lexicalization. The Dutch denominal adjective *natuur-lijk* "natural" has acquired the additional meaning "of course" when used as an adverb, just like English *naturally*. The word has become opaque, and speakers may no longer feel a relationship with the base noun *natuur* "nature". This independence of the highly frequent adverb *natuurlijk* /na:ty:rlək/ manifests itself in the fact that its phonetic form is often reduced to forms like [ty:rlək] or even [ty:k]. This contrasts with the use of this word as an adjective with the regular meaning "natural". Used with that meaning, the word's phonetic form cannot be reduced, and it must be pronounced as [na:ty:rlək].

Words may also lose their morphological transparency because the base word was lost in the course of time. In the English verb *begin*, for instance, we might still recognize a prefix *be-*, but the root *gin* is no longer used as a verb. Hence, we might conclude that this verb has become a simplex one. In some cases there is reason to speak of words that are only formally complex. Such formal complexity may be the result of the loss of the base word, or of borrowing, as happened in English through the influx of Latinate verbs. English verbs in *-ceive* such as *conceive*, *perceive*, and *receive* share the property that their corresponding nominal ends in *-cept-ion*, and their corresponding adjective in *-cept-ive*. For that reason, we would like to assume that these verbs consist of two constituents, a prefix and a root. Yet, there is no systematic form–meaning correspondence involved in such sets of similar words.

20.9 Processing complex words

The processing of complex words is an important domain of psycholinguistic research. The main debate concerns the balance between computation and storage of complex word forms. One position, which has been most eloquently defended by Steve Pinker (Pinker 1999) is that complex words that are irregular are stored in the lexicon, but that regular forms are computed online when the utterance is being processed. For instance, the inflectional forms of English regular verbs are assumed not to be stored, unlike those of the irregular verbs. The different patterns of irregular verb forms may however be recognized and stored in a kind of associative memory.

This view of the balance between storage and computation has been challenged by many psycholinguistic research results. It appears that fully regular forms may induce frequency effects. If a word form has a relatively high frequency of use, this will speed up the processing of that word in processing tasks such as lexical decision tasks. In such tasks subjects have to decide whether a letter sequence shown on a screen is a correct word of the language. The idea behind the frequency effect is that the frequent use of a word heightens its resting activation level (or "lexical strength"), and thus the word is recognized faster. A frequency effect is only possible if the relevant words or word forms are stored, and hence can be subject to frequency effects. Thus, we can conclude that regular word forms may be stored.

These findings can be modeled in a dual route model with competition. When the meaning of a complex word has to be understood, there are two routes available: the complex word is either retrieved directly from lexical memory, or first decomposed into its morphological constituents, on the basis of which the meaning is then computed. Since both routes are available, they will compete. If a complex word is an established one, with a high frequency, and thus with a high resting activation level, the direct route will win. If the word to be understood is a completely new one that is not stored, or has a low frequency of use, the decompositional route will be the fastest (Baayen et al. 1997).

At a more fundamental level, the issue is whether we are justified to make a distinction between symbolic rules or schemas on the one hand, and representations on the other. The distinction between rule and representation is denied in connectionist approaches to language structure. These issues are too complex to be discussed in this chapter; however, see Bybee and McClelland (2005) for a recent general discussion of the issues involved.

Other important results of psycholinguistic research are its findings concerning the structure of the lexicon. It is clear that the lexicon is not a list but rather a network of relationships between words, relations along different dimensions such as phonology, semantics, and morphological structure. The lexicon is a web of words. This means that the paradigmatic relationships between words (either simplex or complex) are essential for understanding how morphology works.

An example of the role of paradigmatic structure is the family size effect (Baayen and Schreuder 2003): a simplex word is more easily processed the larger the number of words derived from it (its morphological family) is. This effect presupposes that a word is linked to its derivatives in the lexicon. Another paradigmatic effect is that the choice of a linking element in Dutch compounds can be very well predicted on the basis of the constituent family of the left and the right constituent. For instance, subjects tend to choose the linking element -*en* for a new compound that begins with the constituent *rat* "rat" since -*en* is the preferred linking element in established compounds that begin with *rat*.

The processing of complex words can also be investigated through naturalistic data such as slips of the tongue, which give a clue as to how complex words may be represented in the mental lexicon. When a complex word is stored, it might be stored including its internal morphological structure. In the following slips of the tongue, two morphemes have been exchanged, which suggests that the internal structure of such words is indeed available (Cohen 1987):

(32) Ik heb vrijdag maan < Ik heb maan-dag vrij
I have Friday moon I have Mon-day free
'I am free on Monday.'
een vloeibare drinkstof < een drink-bare vloeistof
a liquid drink-stuff a drink-able flow-stuff
'a drink-able liquid'

Thus, speech errors can provide linguistic evidence for the way in which complex words are stored in lexical memory.

20.10 Morphological productivity

A much discussed notion in morphological research is the notion of productivity. Morphological processes differ in the extent to which they are used for coining new words or word forms. The classical example from the domain of inflection is the difference between the regular formation of past tense forms of English, which is productive and applies to newly coined verbs, and the class of irregular verbs, which form their past tense by means of vowel change and occasionally consonantal changes as well, a case of internal modification (*sing-sang, bring-brought*, etc.). In the domain of word formation, one may observe that new deadjectival nouns ending in *-th* are hardly ever formed by speakers of English, whereas new deadjectival nouns ending in *-ness* can be made readily: *coolness* is more readily coined than *coolth*.

Quantitative measures of productivity make use of the notion "type frequency". A productive process will result in a large number of different types. For instance, in English the number of types of past tense forms of regular verbs is much higher than the number of types of irregular past tense forms, and the number of types of nouns in *-ness* is much higher than that in *-th*. Baayen makes use of the notion "hapax" in measuring quantitative productivity: the number of types of a particular type of complex word that occur only once in a given corpus (that is, the number of hapaxes) is a good measure of the productivity of the morphological process involved (Baayen 1992).

One approach to this phenomenon of differences in productivity is to consider it as a property of the language system: some morphological processes are unproductive (hence, the relevant set of words cannot be expanded), whereas other processes are productive, and may lead to new words or word forms. The actual, quantitative productivity of a productive process is then determined by two types of factors: the number of potential bases to which the process can apply (system-internal factors), and pragmatic, system-external factors such as the need for a particular word. This is the position taken in Bauer (2001). The number of potential bases depends on the number of linguistic constraints on the morphological process involved: the more constraints there are, the fewer chances that the process can apply and create new forms.

However, Baayen (1992, 2008*a*) and Bybee (1995) have argued that productivity is inherently a gradual notion. Even in the case of processes with a very low productivity, the relevant class of words can occasionally be extended. For instance, one may come across the word *coolth* in language corpora.

As to the non-systemic factors, the productivity of word formation processes may depend on factors such as written vs. spoken language, specific registers, and speech communities. Certain types of word formation are used productively in written language only. The suffix *-ity* is typically used productively in scientific and technical discourse (Baayen 2008*a*).

Processing factors also play a role in productivity. The output of a word formation process is morphologically more transparent and will be readily decomposed in processing if the frequency of the derived word is lower than that of the base word. Decomposition of complex words will strengthen the accessibility of the corresponding morphological schema, and thus increase productivity (Hay and Baayen, 2002).

20.11 Tools for morphological research

A primary source of information about morphology is formed by the descriptive grammars of individual languages which usually give a description of inflection and word formation. The availability of such grammars is essential for typological research. An excellent manual for morphological description is Payne (1997).

Typological databases are mainly based on such descriptive grammars. Among others, the following databases on morphological typology can be found on the internet:

- Surrey Morphology Group: http://www/surrey.ac.uk/LIS/SMG
- Universals Archive of the University of Konstanz: http://typo.uni-konstanz.de/archive/intro
- The Morbo database on compounding, University of Bologna: http://morbo.lingue.unibo.it.

It is important that, in morphological analysis, linguists use the same conventions, and the same glossing rules. The Leipzig glossing rules are used as a standard these days. They can be found on http://www.eva.mpg.de/lingua/resources/glossing-rules.php. Tools for language description are available on www.eva.mpg.de/lingua.

Other tools including the Ethnologue survey of languages can be found on www.sil.org.

Sources for the outputs of word formation are the dictionaries of individual languages. Searching for morphological information has been made much easier thanks to electronic dictionaries, many of them online. However, the role of dictionaries in morphological research is now strongly being reduced in favor of corpora of language use. Corpora do not suffer from the restrictions of dictionaries that the data are filtered by the lexicographer, and are always lagging behind as to what happens in actual language use. In fact, present-day good dictionaries are based on corpora as well. Moreover, corpora provide the possibility to investigate how the productive use of morphological processes correlates with factors of language use and properties of language users. Therefore, corpus-based linguistic research has become indispensable for adequate morphological research (Baayen 2008*a*).

CHAPTER 21

OPTIMALITY THEORY IN PHONOLOGY

MARIA GOUSKOVA

21.1 Introduction

The goal of this chapter is to overview Optimality Theory (OT, Prince and Smolensky 2004) as applied to phonology.[1] OT is a theory of constraint interaction in grammar, which aims to solve a couple of related problems that have confronted generative phonological theory since its earliest days. The first problem is conspiracies: in some languages, there is a constraint that seems to be satisfied in a variety of ways, as if the rules conspire to achieve a single target. The second problem is soft universals: unrelated languages show evidence of the same or similar constraints, but the constraints do not seem to hold in all languages. Moreover, as in single-language conspiracies, the way these constraints are satisfied may differ from

I'd like to think Amanda Dye, Joe Pater, Jen Smith, Jim Wood, and the editors for feedback.

[1] Because of space limitations, the review is necessarily incomplete, but there is no shortage of other article-length treatments of OT (Prince and Smolensky 2002, McCarthy 2007*b*, Tesar et al. 1999, Smolensky et al. 2006) or book-length overviews (McCarthy 2008, 2002, Kager 1999, Archangeli and Langendoen 1997, Prince and Smolensky 2004). In addition to the book-length collections of phonology papers in OT cited in the body of the chapter, there are other general (McCarthy 2003*a*) and topical collections (Lombardi 2001*a*, Roca 1997). Last but not least, there is an extensive free online archive of papers in and about OT, the Rutgers Optimality Archive, at http://roa.rutgers.edu.

language to language. OT addresses both problems by introducing the assumption that constraints are universal but rankable and violable. This simple assumption has many surprising consequences, which have been fruitfully pursued in the fifteen years since the advent of the theory.

Conspiracies were discovered by Kisseberth (1970*a*), who describes several rules in Yawelmani that are united functionally but couldn't be unified formally in the theory of the time (*The Sound Pattern of English*/SPE, Chomsky and Halle 1968). Another example comes from Tonkawa (Kisseberth 1970*b*, McCarthy 1986). Tonkawa has a rule of vowel deletion, which happens to be blocked just in case it would create a geminate, or long consonant (see (1)).

(1) A constraint on vowel deletion in Tonkawa (Hoijer 1933)

a. Vowel deletion between non-identical consonants

/noto̱xo-n-oʔ/	notxonoʔ	'he is hoeing it'	cf.	notox	'hoe'
/pice̱na-n-oʔ/	picnanoʔ	'he is cutting it'	cf.	picen	'castrated steer'

b. No vowel deletion if surrounding consonants are identical

/hewa̱wa-n-oʔ/	hewa̱wanoʔ	'he is dead'	*hewwanoʔ
/ham'a̱m'a-n-oʔ/	ham'a̱m'anoʔ	'he is burning'	*ham'm'anoʔ

Another rule of vowel deletion, which deletes the stem-final vowel in compound formation, may apply even between identical consonants, but what surfaces is a single short consonant (see (2)). When two identical consonants are brought together by morpheme concatenation, one of them also deletes (see (3)). In this conspiracy, the rules of vowel deletion and consonant shortening work together to avoid geminates:

(2) Tonkawa compound vowel deletion and geminate simplification (McCarthy 1986: 225)

/taʔane̱-nosʔo:yta-/	taʔanosʔo:yta-	'to stretch (e.g., a rope)'
/yakona̱-nacaka-/	yakonacaka-	'to kill (him) with a blow of fist'
/yakexe̱-xakana-/	yakexakana-	'to push (it) down hard'

(3) Tonkawa morpheme concatenation and geminate simplification (Hoijer 1949)

/nes-so:la-/	neso:la-	'to spill (it)'	cf.	nes-kapa-	'to shut a door'
/ʔey-yace-/	ʔeyace-	'to catch, capture (them)'	cf.	ʔey-pake-	'to slice it'

This anti-geminate prohibition applies not only to derived geminates: there are no geminates even morpheme-internally in the language, so morphemes like **hewwa*- are absent.

(4) A constraint on Tonkawa morphemes

$^{*} \ldots C_i C_j \ldots$, where i = j

An insightful analysis of Tonkawa would explain why the geminate prohibition holds both in derived and in underived sequences. It should also capture the obvious connection between this prohibition blocking vowel deletion and triggering consonant shortening. Yet pre-OT treatments had to explain such generalizations through two separate mechanisms. Restrictions on underived sequences were handled through Morpheme Structure Constraints, which held at the underlying level. Restrictions on derivations were stated as part of the rule's context, or else put into a separate "derivational constraint" whose interaction with the rule it blocked was never fully formalized (Kisseberth 1970*b*). It was likewise impossible to draw any connection between such constraints and the rules they trigger, as in the case of consonant shortening. Any similarity between a Morpheme Structure Constraint and a condition on some rule's application was purely coincidental, creating a redundancy known as the *duplication problem* (see McCarthy 2002, §2.1 for discussion). OT avoids the duplication problem by assuming that the constraint against geminates applies only to surface forms, prohibiting both derived and underlying geminates. It blocks vowel deletion because it overrides the constraints that make vowel deletion necessary in the first place. OT can also make sense of the way the geminate prohibition apparently triggers consonant shortening because constraints in OT can interact to compel such alternations.

The prohibition against geminates is a kind of soft universal. It blocks vowel deletion in many unrelated languages such as Afar and Tiberian Hebrew (McCarthy 1986). As Odden (1988) shows, however, there are languages where this is not the case.[2] Even in languages that freely violate the constraint against geminates, there is often evidence that they are disfavored. In a theory without violable constraints, counterexamples to a purportedly universal constraint immediately put its validity into question. Existing solutions are all somewhat unsatisfying—for example, sometimes it is posited that the principles hold at different levels of derivation in different languages, or they are treated as parameters with language-specific settings. Nevertheless, theories with inviolable constraints have no way of capturing the intuition that the same constraint seems to be at work even though it appears to be violable. On the other hand, soft universals are unsurprising for OT, since OT constraints are violable and universal: languages may either satisfy constraints or skirt them altogether. OT furthermore predicts that constraints can be satisfied partially even if they are generally violated in a language. This kind of interaction, of which there is ample evidence in work on OT, is known as the Emergence of the Unmarked (McCarthy and Prince 1994).

By assuming that constraints are universal and violable, OT suggests a natural theory of typology. In the strongest version of the theory, constraints are universal, and any reordering of them should produce a plausible grammar. This simple premise makes for an easily falsifiable theory of phonological grammar, and it has many interesting consequences. For example, OT allows for a principled approach

[2] There are still others that have an almost reverse pattern of deleting a vowel only between identical consonants; for a recent OT account and typology, see Baković (2005).

to the problem of variation. Variation can be framed as two or more different grammars that coexist within a speaker or a community. Since OT explicitly formalizes the notion of differences between grammars, it can account for variation with a few modest extensions. The problem of learning can also be understood in similar terms: how does a learner arrive at the right grammar when starting out with an incorrect one?

The chapter is organized as follows. Section 21.2 describes the architecture of the theory, including its basic components (§21.2.1–21.2.3.2) and approach to typology (§21.2.3.3). Section 21.3 addresses the status of the lexicon in OT. Sections 21.4 and 21.5 describe some work on learnability, acquisition, and variation. Section 21.6 concludes.

21.2 Architecture

An OT grammar has three components. Con is the component that defines the set of universal violable constraints. Gen is the component where output candidate parses are generated based on input forms. Eval is the component that selects an optimal output from the set of alternative candidates, given a language-specific hierarchical ordering of Con, $\mathcal{H}$. The path from the input to the output is charted in (5). Even though versions of OT may differ from each other in the details of how Gen and Eval are implemented, most work in OT assumes something like (5).

(5) OT: the organization of the grammar

/input/	→	Gen	→	Eval($\mathcal{H}_{CON}$)	→	[output]
		candidate 1				
		candidate 2				
		candidate 3				
		...				

Each of these components is examined in turn in the following sections, starting with Eval (§21.2.1) and moving on to Gen (§21.2.2) and Con (§21.2.3.2). The focus throughout will be on phonological issues and applications, set in the most widely accepted version of the theory known as "classic" or "parallel" OT (Prince and Smolensky 2004, McCarthy and Prince 1995).[3] The main principles of this version of the theory are outlined in Prince and Smolensky (2004); since it is impossible to do justice to all the work done on the theory since 1993, interested readers are referred to the various works cited along the way for more recent developments.

[3] Several proposals depart from this architecture by assuming a constrained or modified Gen (McCarthy 2007*a*) or multiple serially ordered evaluations by the constraint hierarchy (Stratal OT; Kiparsky to appear). Some versions include an additional component that further filters the output of Eval (Orgun and Sprouse 1999, de Lacy 2007).

21.2.1 Eval

One of the defining features of OT is competition between candidates, and this competition is resolved in the Eval module. Eval considers candidates in pairwise comparisons based on their relative performance with respect to the constraint hierarchy. For an informal example, consider the Tonkawa pattern described above. Given the input /hewawa-n-oʔ/, the grammar has a choice between a candidate that deletes the vowel, *[hewwanoʔ], and one that does not—the actually attested [hewawanoʔ]. Vowel deletion is required by the general phonology of Tonkawa, so not deleting violates a constraint, but deleting creates a geminate, which also violates a constraint. The fact that deletion is blocked suggests that "NoGeminates" dominates "DeleteVowel". (These constraints are just placeholders for now; I return to the actual constraints at work in §21.2.3.2.)

More abstractly, two constraints disagree on a pair of candidates when one of the constraints favors the candidate that the other constraint disprefers. In the schematic example (6), Constraint 1 and Constraint 2 disagree in just that way on outputs 1 and 2 (an asterisk in a constraint's column indicates that the candidate violates it). Output 1 is more *harmonic* than output 2 with respect to Constraint 1, and the opposite is true for Constraint 2. We infer that Constraint 1 dominates Constraint 2 because output 1 is the surface form. Its *optimal* status is marked here with ☞:

(6) Candidate evaluation by constraints

/input/	Constraint 1	Constraint 2
a.☞ output 1		*
b. output 2	*!	

Proof of constraint ranking is given in a *tableau* such as (6), or in the more compact format of a comparative tableau (see (7)). In a comparative tableau, the optimum (given first) is paired with a loser. Since, as shown above, Constraint 1 prefers the winner of the competition, a W appears in its column. An L in the column of Constraint 2 indicates that it prefers the loser.

(7) A comparative tableau

/input/	Constraint 1	Constraint 2
output 1_W~output 2_L	W	L

The optimum is the candidate that best satisfies the highest-ranked constraint that distinguishes it from other candidates. It is often the case that for a given winner~loser pair, some constraint does not distinguish the candidates, either because both violate the constraint or because both satisfy it to the same extent. In such cases, the constraint does not participate in determining the outcome, and the optimum must be chosen by some other constraint. Thus, if each of the outputs in (7) violated some Constraint 3 three times, it would not affect the optimality of output 1 even if that constraint happened to be ranked above Constraint 1. What matters is not absolute satisfaction of constraints but rather comparative performance.

This model of Eval is assumed in most OT work, both in phonology and in other areas (though there are other theories of Eval, e.g., Wilson 2001). Notably, this usually ensures a unique winner for any competition, since there is usually at least some constraint that distinguishes even very similar candidates. This architecture of Eval must be changed, however, in order for the theory to produce variation or optionality; some proposed modifications are discussed in §21.5.

21.2.2 Gen

Gen is the component of an OT grammar that generates the competing candidates from which the output is chosen. Since there is no primitive notion of rules or transformations in OT, it falls to Gen to produce a wide enough range of forms that would cover the range of phonological operations, though it even goes beyond that, as we will see shortly. Basically, the job of Gen is to improvise on the input. In phonology, Gen can map an input to an output more or less without changes, or it can modify the input. There are several ways to render an input faithfully. An input like /patra/ can be syllabified as either [pa.tra] or [pat.ra]. Gen can also manipulate the input by changing distinctive features, deleting or inserting segments, assigning stress, and so on. In addition to assigning phonological structure, Gen tracks how each input is mapped to each output candidate by positing a relation between elements of the input and the elements of the output. Thus, each candidate is not just an output form but also a mapping from the input. This input–output relation is typically formalized in Correspondence Theory (McCarthy and Prince 1995). Some examples of candidates for a hypothetical input /bak/ are given below.

(8) Some candidates emitted by the phonological Gen

input	*output*	*comments*
/bak/	bák	fully faithful candidate
	bá.ki̲	epenthesis, stress on the first syllable
	ba.kí̲	epenthesis, stress on the second syllable
	bá	deletion
	i̲.bá.ki̲	double epenthesis
	káb	metathesis (reordering)
	vák	feature change (frication)
	pák	feature change (devoicing)
	...	

An important property of GEN in traditional OT is that it is unconstrained by phonological well-formedness principles and knows nothing about constraint satisfaction. Instead, GEN is usually assumed to have *Freedom of Analysis* (McCarthy and Prince 1993*b*): it can modify the input in all sorts of ways, many of which may seem absurd to a linguist. For example, given the hypothetical input /pata/, GEN is allowed to map it to the sensible [(pá.ta)] or [(pa.tá)],[4] but also to [(pá)(tà)], [(pá.ta)ʔə], [(pán.da)], [(á.tap)], [(p(á)t.a.)], and so on. This somewhat counter-intuitive feature of OT stems from the imperative of decoupling grammatical operations from the constraints that determine what surfaces. The operations themselves are of relatively little interest, but well-formedness is paramount, and it is a matter for constraints to sort out.

This rich array of candidates emitted by GEN has been argued to be problematic. At least in part, it contributes to the so-called too-many-solutions problem (a term due to Steriade 2001): there are many conceivable ways of avoiding certain marked structures, yet many seem to be unattested cross-linguistically. One simple example is avoidance of final voiced obstruents. While word-final devoicing (/pad/ → [pat]) is extremely common, vowel epenthesis (/pad/ → [pa.di]) and consonant deletion (/pad/ → [pa]) appear to be unattested. This is all the more puzzling since laryngeal features are different in this respect from place features. Lombardi (2001*b*) proposes a solution that relies in part on a revision of GEN and in part on certain assumptions about CON, but similar problems arise in other areas, so there is no universal solution. It should be noted that the too-many-solutions problem is an issue not just for OT but for any theory of phonology that aims to account for typology: if the right theory of phonology is rule-based, there is still a question of why certain rules seem to be ubiquitous and others don't seem to occur, which was never satisfactorily solved.

Some recent work, however, challenges Freedom of Analysis (see the various contributions to Blaho et al. 2007). In McCarthy's (2007*a*) OT with Candidate Chains, candidates are generated in incremental steps, and each step is checked against the constraint hierarchy to ensure that the change increases well-formedness. McCarthy argues that this revision can address the too-many-solutions problem; see also Wilson (2001). Interesting results may also be obtained by changing the way GEN manipulates segmental, syllabic, and metrical structure, and there are arguments that this is actually a necessary restriction on the theory (Morén 2007, Smith 2007, Rice 2007, Lombardi 2001*b*).

A typical example of what could be at stake comes from the realm of metrical foot structure, discussed by Rice (2007). Most modern work on metrical stress theory assumes that feet are binary pairings of weak and strong elements, drawn from syllables or moras (Prince 1985, Baković 1998). Other theories admit feet

[4] Square brackets signify prosodic words, round brackets delimit feet, and dots show boundaries between syllables. Acute accents stand for primary word stress, and grave ones for secondary stress.

with more than two syllables (Kenstowicz 1996, Hayes 1995). There is one area where such feet could offer an analytical advantage: stress in languages such as Cayuvava, which falls on every third syllable (see Halle and Vergnaud 1987, Levin 1988, and cf. Elenbaas and Kager 1999). As Rice points out, any theory of ternary stress couched in OT must confront two separate questions. The first is whether ternary feet are necessary for analyzing ternary stress. The answer is no—analyses in terms of binary feet are possible. The second question is whether feet of three syllables and larger should be ruled out universally, and if so, how. Some properties of feet can be derived through constraint interaction alone, but others may need to be stipulated to be properties of Gen. If ternary feet are ruled out in Gen, there needs to be a principled theory behind such a prohibition, which is at present lacking.

OT is in principle compatible with different assumptions about phonological structure, since it is a theory of the architecture of the grammar rather than of phonological representations. It is, however, impossible to discuss properties of Gen without making specific theoretical assumptions about substantive properties of phonological theory, i.e., feet, syllables, and features. This is an area of ongoing and future work.

One final issue that is relevant here is the lack of derivations in parallel OT. Unlike SPE and much work in syntactic theory in the generative tradition, OT has just a single-step mapping from the input to the output instead of incremental derivational steps. This addresses certain problems such as top-down interactions between different levels of structure (see Prince and Smolensky (2004) on Tongan, for example), but it introduces another problem, namely a difficulty with certain types of opaque interactions. The problem is too complex to review here in any detail, but there are several proposed solutions. They include reintroducing derivational levels (Kiparsky to appear), special candidates that mimic derivational stages (McCarthy 2003*c*), and, finally and most relevantly, a ground-up revision of Gen that actually includes whole derivations as candidates (McCarthy 2007*a*). The latter book includes a comprehensive overview of the issue and work both in OT and in other theories.

21.2.3 Con and Factorial Typology

In this section, I overview some aspects of the constraint component Con that are assumed in much modern work in phonology. The prevailing theme here is that Con is not homogeneous, and it is not an arbitrary list of ad hoc constraints; rather, it has elaborate structure. At the very least, constraints are classified into markedness, faithfulness, and interface constraints. Within each of these types, there are subfamililes, grouped based on the way they relate to linguistic primitives. I discuss each type in turn. The last subsection deals with OT's approach to typology.

21.2.3.1 *Internal structure of CON: constraint types*

Intuitively, markedness constraints ban elements that are structurally complex or in some sense difficult or disfavored. For example, in the realm of syllable structure, complex onsets such as [pra] are marked compared to simplex onsets such as [pa]; hence, there is a markedness constraint *Complex against tautosyllabic consonant clusters. The definitional property of markedness constraints, however, is not that they ban difficult things but rather that they refer only to output structures. Markedness constraints are often but not always based on phonetic principles; they may also have formal origins. Constraints of both types are discussed in the following sections.

By contrast, faithfulness constraints govern disparities between two levels of representation. The most familiar levels are the input (underlying representation) and the output (surface representation); for example, the constraint Max (McCarthy and Prince 1995) requires every segment in the input to have a correspondent in the output. This requirement is violated by deletion (e.g., /pra/ → [pa]). Almost every kind of disparity[5] between input and output violates some faithfulness constraint: deletion, insertion, reordering/metathesis, featural changes, and other operations of Gen discussed in §21.2.2 all have associated faithfulness costs. Faithfulness constraints in phonology can mediate between segmental strings at other levels of representation, as well. Since the same kinds of disparity sometimes result from inexact copying in the domain of morphological reduplication, McCarthy and Prince (1995) propose a unified theory of faithfulness, according to which the same types of constraints mediate between input vs. output and reduplicative base vs. reduplicant copy. Benua (1997) extends the theory to apply between words that are related by morphological derivation, and other proposals have since extended faithfulness to other domains.

To illustrate the markedness–faithfulness distinction, consider Tonkawa vowel deletion. Recall that in Tonkawa, vowels delete between two non-identical consonants but not between identical ones. The first step in analyzing Tonkawa is explaining why vowel deletion happens at all. Deletion creates a mismatch between the input and the output, which violates faithfulness. This means that some markedness constraint dominates faithfulness. The logic here is intuitively simple: given the way Eval works, there must be a tradeoff for violating faithfulness, and the only reason for a candidate to map unfaithfully is to become less marked.[6] Suppose the tradeoff is satisfying the requirement that stressed syllables be long, or heavy: if the vowel is deleted, the remaining consonant can close the preceding syllable.[7]

[5] Not all operations of Gen do, though. It is usually assumed that syllabification is not associated with faithfulness violations, since it is not contrastive. If a faithfulness constraint were violated by positing or changing syllable structure, then we would expect to see contrasts in syllable structure alone. See, for example, McCarthy 2003*c*.

[6] This property of OT is known as Harmonic Ascent (McCarthy 2002, Moreton 2003).

[7] A full analysis of Tonkawa vowel deletion along these lines is developed in Gouskova (2003).

This is shown in (9): the winner is unfaithful, since it deletes the second underlying vowel, and the loser is marked, since it has a stressed open syllable [nó...]. The markedness constraint Stressed=Heavy, which requires stressed syllables to be heavy, dominates the faithfulness constraint Max-V, which prohibits the deletion of vowels:

(9) Tonkawa: vowel deletion required between non-identical consonants

/notoxo-n-oʔ/	Stressed=Heavy	Max-V
nót.xo.nóʔ∼nó.to̲.xo.nóʔ	W	L

Markedness and faithfulness constraints are often in conflict, but markedness constraints can also conflict with other markedness constraints. In Tonkawa, the constraint against geminates, NoGem, blocks the normal application of vowel deletion. This conflict is shown in (10). Here, the markedness constraints conflict with each other since the choice is between having a geminate on the one hand and having a stressed open syllable on the other. No-Gem actually agrees with Max-V on the candidates in 10. Thus, whether markedness and faithfulness constraints conflict really depends on the candidate and the constraints in question.[8]

(10) Tonkawa: vowel deletion blocked between identical consonants

/hewawa-n-oʔ/	NoGem	Stressed=Heavy	Max-V
hé.wa̲.wá.noʔ∼héw.wa.nóʔ	W	L	W

The third type of constraints in OT are interface constraints generated by Generalized Alignment (discussed in §21.2.3.2.1) and morpheme realization constraints, which require that morphological entities be realized as phonological content (see McCarthy and Wolf 2005 for recent discussion). These share some features with faithfulness, since they also mediate between two levels of structure, but instead of looking at the same type of structure instantiated in the input and the output, they require access to structures from different components of the grammar. A typical interface constraint will require a given edge of a phonological structure to coincide with some edge of some morphological structure. An example would be the requirement for a phonological word to end with a segment that belongs to a lexical/morphological word (Selkirk 1995, McCarthy and Prince 1993*a*). This requirement is violated by enclitics, which are prosodified with the preceding lexical word (e.g., English possessive *Mary's̲*), and by inserted segments, which are assumed not to have any morphological affiliation at all (e.g., Lardil augmentation /kaŋ/ → [kaŋk̲a̲], Prince and Smolensky 2004). Unlike the markedness/faithfulness dichotomy, such interface constraints are not an essential feature of the

[8] This analysis of Tonkawa vowel deletion has not yet addressed the problem brought up at the outset, namely the general absence of tautomorphemic geminates in Tonkawa, but I return to it in §21.5.

architecture of OT; indeed, they are very much a feature of a particular theory of the interface between morphosyntax and phonology (Selkirk and Shen 1990 and others).

Thus far, we've seen three types of constraints: markedness, faithfulness, and interface constraints. They are classified based on the structural levels they access in the process of evaluation. Constraints may further be classified based on the way they relate to linguistic primitives such as phonetic scales and atomic elements of phonological representation; this is the subject of the next section.

21.2.3.2 *Internal structure of CON: constraint schemata*

The content of the constraint component of the phonological grammar should be the most controversial aspect of the theory, since its typological predictions depend on how the constraints interact with each other under re-ranking. Unsurprisingly, a considerable effort in early OT work was devoted to discovering the constraints and working out their relationships to each other and to the substantive principles thought to underlie phonological patterns. One of the most productive lines of attack on this has been the development of constraint schemata, which define families of constraints based on how they are built from phonological primitives. The two best-known constraint schemata are Generalized Alignment (McCarthy and Prince 1993*a*) and Harmonic Alignment (Prince and Smolensky 2004). I discuss each in turn.

21.2.3.2.1 *Generalized Alignment* Generalized Alignment is a constraint schema proposed by McCarthy and Prince (1993*a*) in the context of a theory of edge effects. Constituent edges are often the domain of special phonology. Thus, stress is attracted to word edges: descriptions of stress patterns often make references to initial, final, and penultimate syllables. Which edge wins as a default is up to the language, but the orientation of stress toward word edges is a near-universal feature of stress (Hayes 1995). Similarly, languages may differ in whether they require prosodic word edges to coincide with lexical word edges or whether there can be mismatches between the two types of structures. Under Generalized Alignment, these sorts of observations are captured by constraints that require edges of particular domains to coincide with edges of other domains: metrical feet must be aligned to prosodic word edges, lexical words must be aligned to prosodic word edges, and so on. The basic ingredients of an alignment constraint are constituents to be aligned and the edge(s) that must coincide. McCarthy and Prince (1993*a*) define Generalized Alignment as follows:

(11) Generalized Alignment
Align (Cat1, Edge1, Cat2, Edge2)$=_{\text{def}}$

∀ Cat1 ∃ Cat2 such that Edge1 of Cat1 and Edge2 of Cat2 coincide.
Where Cat1, Cat2 ∈ PCat ∪ GCat
Edge1, Edge2 ∈{Right, Left}

In the original proposal, Generalized Alignment applied to prosodic (Selkirk 1978 and others) and morphological constituents, as shown in (12). It has since been extended to many other phonological structures and representational primitives, including subsegmental features (Kirchner 1993), tones (Myers 1994), metrical grids (Gordon 1999). Alignment is such a general formalism for constraint definitions that some have proposed to rethink even familiar constraints such as ONSET and NOCODA in alignment terms: ONSET requires simply that the left edge of a syllable must coincide with the left edge of a consonantal segment (McCarthy and Prince 1993*a*, Ito and Mester 1994). Similarly, Ito and Mester (1994) propose that the old CODACOND of Ito (1986) should be undersood as a family of alignment constraints that require certain features that are marked in coda position to be aligned to the beginning of a syllable rather than the end.

(12) Categories referenced by alignment

PCat = levels in the Prosodic Hierarchy	GCat = morphological constituents
ProsodicWord	MorphWord → Stem*
Foot	Stem → Stem, Affix
syllable	Stem → Root

An issue separate from which constituents to align is how edge alignment constraints assign violation marks. This is actually a problem that extends beyond alignment to all constraint evaluation. In the original proposals (Prince and Smolensky 2004, McCarthy and Prince 1993*a*), some alignment constraints were assumed to be violated gradiently: a single instance of a misaligned structure could incur more than one violation of the constraint, depending on degree of deviation from perfect alignment. McCarthy and Prince (1993*a*) assume that metrical foot alignment constraints such as ALL-FEET-RIGHT "The right edge of each foot corresponds with the right edge of some prosodic word" assign a violation mark for each syllable that stands between the foot and the prosodic word edge. If there is more than one foot in a word, each foot's misalignment contributes to total violations:

(13) Gradient evaluation of edge alignment

	ALL-FEET-RIGHT	*comments*
a. $\sigma\sigma(\acute{\sigma}\sigma)$	✓	perfect alignment on the right
b. $\sigma(\acute{\sigma}\sigma)\sigma$	*	one foot, misaligned by one syllable
c. $(\acute{\sigma}\sigma)\sigma\sigma$	**	one foot, misaligned by two syllables
d. $(\acute{\sigma}\sigma)(\acute{\sigma}\sigma)\sigma$	****	two feet, misaligned by one and three syllables

Gradient evaluation captures certain aspects of edge-sensitive phenomena very well: for example, it offers a straightforward analysis of antepenultimate stress (as in Macedonian, for example) as the resolution of a conflict between NonFinality and All-Feet-Right. Intuitively, in Macedonian, the foot is placed as close as possible to the right edge of the word, but not so close as to encompass the last syllable; since perfect right-alignment is impossible with NonFinality(Foot) ranked above All-Feet-Right, the next best option is chosen instead because alignment is violated only minimally. Unfortunately, this approach also appears to overgenerate by predicting certain unattested patterns in stress, infixation, and harmony systems (Kager 2001, McCarthy 2003*b*). This overgeneration is one of several arguments (McCarthy 2003*b*) for the claim that all constraints are categorical: a candidate should only incur multiple violations of a constraint if it has more than one instance of a structure that violates the constraint. As noted earlier, however, this controversy is separate from Generalized Alignment as a substantive theory of edge phonology.

21.2.3.2.2 *Harmonic Alignment* Harmonic Alignment is a theory of how phonetic and other extralinguistic scales are expressed in the grammar. In the most general form, Harmonic Alignment postulates that there is a relation between prominence and position: prominent positions are ideally filled with prominent elements, and non-prominent positions are filled with non-prominent ones. By now, Harmonic Alignment has been productively extended to generate phonological constraints on sonority-sensitive stress (Kenstowicz 1996), positional vowel reduction (Crosswhite 1999), and tone-stress interactions (de Lacy 2002*a*). Originally, however, Prince and Smolensky (2004) proposed Harmonic Alignment specifically to capture the well-known role of sonority in syllabification (see, for example, Clements 1990), and I'll discuss this application of it here. Prince and Smolensky observe that the prominent position of syllable nucleus is ideally filled with the most sonorous segment, i.e., a vowel, and the non-prominent position of syllable margin (i.e., onset) is ideally filled with an obstruent. Under their proposal for Harmonic Alignment, the position and prominence scales in (14) and (15) would combine to give two scales, one of which defines harmony for nuclei, and the other for margins. Harmonic Alignment is formulated as follows:

(14) Syllable position
Nucleus > Margin (Onset)

(15) Sonority scale
Vowels > Liquids > Nasals > Obstruents

(16) Harmonic Alignment (Prince and Smolensky 2004)
Given binary dimension D_1 with a scale $X > Y$ on its elements $\{X, Y\}$, and another dimension D_2 with a scale $a > b > \ldots > z$ on its elements, the *harmonic alignment* of D_1 and D_2 is the pair of harmony scales:
H_x: $X/a \succ X/b \succ \ldots X/z$ ["$\succ$" means "is more harmonic than"]

H$_Y$: Y/z≻ ... Y/b≻Y/a
The *constraint alignment* is the pair of hierarchies:
*X/z≫ ... ≫*X/b≫*X/a
*Y/a≫*Y/b≫ ... ≫*Y/z

Combining the scales in (14) and (15) gives us the following constraint hierarchies:[9]

(17) *Onset sonority:* *Ons/Vowel≫*Ons/Liquid≫*Ons/Nasal≫*Ons/ Obstruent

(18) *Nucleus sonority:* *Nuc/Obstruent≫*Nuc/Nasal≫*Nuc/Liquid≫ *Nuc/Vowel

This pair of hierarchies, as others in Harmonic Alignment theory, has a special status in Con. Whereas normally, the rankings of constraints may freely vary from language to language (see §21.2.3.3), the relative ranking of the constraints above *with respect to each other* is universally fixed. They can be interspersed with other constraints, so they do not need to be adjacent in a specific language's hierarchy, but it can never be the case that *Nuc/Liquid dominates *Nuc/Nasal, for example, making syllabic liquids more marked than nasal ones.

This explains a well-established typological property of syllabification (Bell 1978): if less sonorous segments can be nuclei in a language, then the language must also allow all the more sonorous segments to be nuclei. Thus, as shown in (19), in some languages, only the most sonorous segments such as vowels may serve as syllable nuclei, whereas in others syllable nuclei can include vowels and liquids, and in still others—vowels, liquids, and nasals. No language allows nasal syllable nuclei without also admitting liquid and vocalic ones, all else being equal. The converse holds for syllable margins (Clements 1990, Steriade 1988, Hankamer and Aissen 1974).

(19) Sonority of syllable nuclei: a typology[10]

	Vowels	Liquids	Nasals	Obstruents
Spanish, Russian	✓	✗	✗	✗
Macedonian, Czech	✓	✓	✗	✗
English (unstressed syllables), Setswana	✓	✓	✓	✗
Berber, Central Carrier	✓	✓	✓	✓
Unattested	✗	✗	✓	✓

[9] Prince and Smolensky call their constraints *P/x and *M/x for "peak" (= nucleus) and "margin" (= onset), respectively, and their sonority scale includes more detail—which I abstract away from here.

[10] Language sources: Spanish (Harris 1983), Russian and Czech (Townsend and Janda 1996), Macedonian (Crosswhite 2001), English (Borowsky 1986), Setswana (Coetzee 2001), Berber (Prince and Smolensky 2004, Dell and Elmedlaoui 1985), Central Carrier (Walker 1979).

Fixed rankings are designed to explain this typological observation. Any constraint that dominates *Nuc/Nasal, for example, will have to dominate *Nuc/Liquid and *Nuc/Vowel, since they are universally ranked below *Nuc/Nasal. A language with this ranking will tolerate vocalic, liquid, and nasal nuclei rather than violate the relevant constraint. For a concrete example, consider Standard American English. In English, a word-final two-consonant cluster is syllabified into a separate syllable if the second consonant is a sonorant, but not if it is an obstruent.

(20) English syllabification (Borowsky 1986, Levin 1985)

syllabic liquids		syllabic nasals		no syllabic obstruents	
peɪ.pɹ̩	'paper'	bʌ.ʔn̩	'button'	mɪks (*mɪ.ks̩)	'mix'
bʌ.kl̩	'buckle'	rɪ.ðm̩	'rhythm'	æsk (*a.sk̩)	'ask'

The conflict here is between creating a tautosyllabic consonant cluster, which violates the markedness constraint *Complex, and tolerating a consonantal nucleus, which violates one of the *Nuc/x constraints. In English, the solution is to tolerate sonorant consonantal nuclei but not obstruent ones, which suggests the ranking in (21). Syllabification patterns in English are thus non-uniform, with a cut-off point at nasals for minimum nucleus sonority.[11]

(21) Syllabification in English: sonorant but not obstruent consonants are parsed as nuclei

	*Nuc/Obs	***Complex**	*Nuc/Nas	*Nuc/Liq	*Nuc/Vowel
peɪ.pɹ̩~peɪpɹ		W		L	
bʌ.ʔn̩~*bʌ*ʔn		W	L		
mɪks~mɪ.ks̩	W	L			

Compare the English pattern with that of Russian. Russian has only vocalic nuclei and freely tolerates margin clusters. The ranking of *Complex with respect to the *Nuc/x hierarchy in Russian must be as in (23):

(22) Russian syllabification: no syllabic consonants at all

metr	'meter'	dogm	'dogma Gen. Pl.'
so.fokl	'Sophocles'	fe.niks	'phoenix'

(23) Russian syllabification: tautosyllabic clusters are chosen over consonantal nuclei

	*Nuc/Obs	*Nuc/Nas	*Nuc/Liq	*Complex	*Nuc/Vowel
metr~me.tr̩			W	L	
dogm~do.gm̩		W		L	
fe.niks~fe.ni.ks̩	W			L	

[11] An additional complication in English syllabification (in unstressed syllables) is that consonants may not by syllabic after a segment of greater sonority; thus, we get [fʌ.nl̩] "funnel" but [kɪln] "kiln". Relational constraints on sequences are also derived by schemata building on Harmonic Alignment; see Baertsch (2002), Gouskova (2004).

To complete this typology, consider the other rankings of *Complex with respect to the *Nuc/x hierarchy. If *Complex is ranked between *Nuc/Nasal and *Nuc/Liquid, the resulting grammar allows only liquid and vowel nuclei, whereas nasals and obstruents will be syllabified into clusters. This is what we find in a number of Slavic languages such as Czech and Macedonian. In Imdlawn Tashlhiyt Berber, on the other hand, any consonant may serve as a syllable nucleus, so sequences of consonants are syllabified into their own syllables rather than into margin clusters. The resulting typology is shown in (24):

(24) Factorial typology of *Nuc/x and *Complex

***Complex**≫*Nuc/O≫*Nuc/N≫ *Nuc/L≫*Nuc/V	any segment can be syllabic	Berber
*Nuc/O≫***Complex**≫*Nuc/N≫ *Nuc/L≫*Nuc/V	syllabic sonorants, but not obstruents	English
*Nuc/O≫*Nuc/N≫***Complex**≫ *Nuc/L≫*Nuc/V	syllabic approximants	Czech
*Nuc/O≫*Nuc/N≫*Nuc/L≫ ***Complex**≫*Nuc/V	only vowels can be syllabic	Russian

By now it should be apparent why the hierarchy of *Nuc/x constraints must be fixed. If *Nuc/x constraints could be reranked with respect to each other, then the theory would not make any predictions regarding typological implicational universals. If the ranking Nuc/Liq≫*Complex≫*Nuc/Obs were possible, we would expect to see languages that have syllabic obstruents but not liquids. Such languages are unattested, so the possibility of such a ranking must be excluded.

The usual motivation for fixed rankings is that they reflect extragrammatical principles. The reason the *Nuc/x hierarchy is fixed is that it is based on the sonority scale, which reflects physical properties of sounds (such as intensity; see Parker 2008 for a recent overview). The strongest version of OT would only admit externally motivated universally fixed rankings, since fixed rankings are a kind of stipulation.[12] Indeed, there are plenty of proposals for universal constraint hierarchies that are not generated by Harmonic Alignment but are still based on phonetic and perceptual scales (Kirchner 1998, Flemming 1995, Steriade 2001, Kawahara 2006, de Lacy 2002*a*). There is a broad consensus in the literature on OT that many phonological markedness constraints are substantively grounded. Whether it is possible to reduce all phonological constraints to primitives, however, is a subject of ongoing work (Hayes 1999, Hayes et al. 2004, Smith 2002).

[12] Another approach to hierarchies is to formulate constraints so that no matter how they are ranked, the universally most marked structures such as syllabic obstruents remain more marked than syllabic sonorants (Prince 1998, de Lacy 2004).

21.2.3.3 *Typology in OT*

As other generative theories, OT aims not only to delimit the range of cross-linguistic variation but also to derive universals. Both of these questions are addressed through a single mechanism: constraint re-ranking. OT offers a novel and strong hypothesis regarding cross-linguistic typology. According to the hypothesis, the range of cross-linguistic variation is determined by the number of constraint rankings that yield distinct sets of surface forms and mappings. The number of distinct rankings of constraints is the factorial[13] of the cardinality of CON, n!. Since factorials get quite large as n increases (e.g., 6! = 720, but 8! = 40,320), it is essential to demonstrate that OT does not overgenerate distinct grammars. For most realistic constraint sets, there are far more constraint rankings than there are distinct outcomes, since many constraints do not conflict with each other, and others only conflict when dominated by other constraints. To take a simple example, consider a miniature model of CON below, which consists of three constraints discussed in the Tonkawa example (MAX-V, NOGEM, and STRESSED=HEAVY). This constraint set has 3! = 6 permutations, but only three distinct outcomes: languages with no vowel deletion around stressed syllables (such as Spanish), languages with deletion blocked between identical consonants (Tonkawa), and languages with deletion regardless of consonantal context (Klamath, according to Baković 2005).

(25) A mini-CON and factorial typology

Rankings	sample mappings	Pattern and language
MAX-V≫S=H≫NOGEM NOGEM≫MAX-V≫S=H MAX-V≫NOGEM≫S=H	/pataka/→(pá.ta)ka /patata/→(pá.ta)ta	Vowel does not delete (Spanish)
NOGEM≫S=H≫MAX-V	/pataka/→(pát)ka /patata/→(pá.ta)ta	Vowel deletes except between identical Cs (Tonkawa)
S=H≫NOGEM≫MAX-V S=H≫MAX-V≫NOGEM	/pataka/→(pát)ka /patata/→(pát)ta	Vowel deletes regardless of context (Klamath)

There are only three distinct outcomes in this typology, even though there are six rankings. The reason is that MAX-V and NOGEM do not really interact with each other. If MAX-V dominates STRESSED=HEAVY, vowel deletion is not an option, so it doesn't matter where NOGEM is ranked. If STRESSED=HEAVY dominates MAX-V, on the other hand, then all that matters is the relative ranking of STRESSED=HEAVY and NOGEM. Depending on the ranking, deletion will either

[13] "Factorial" ($n!$) is the number of permutations of n elements, which is 1 if $n = 0$ and n*factorial($n-1$) if $n > 0$. The factorial is calculated by multiplying all numbers m: $0 < m \leq n$. Thus, the factorial of 2 is 1*2 = 2, the factorial of 3 is 1*2*3 = 6, and so on.

be blocked or not, but two of the rankings amount to the same outcome. The more complex the constraint set, the richer the possibilities for such interactions, of course, and the number of possible distinct grammars cannot be predicted simply from the number of constraints. The nature of the constraints in question is crucial to working out the typology of possible phonological systems. For this reason, factorial typology is the primary means of testing proposed constraints.

The second typological concern is addressing universals: structures and patterns that either occur in all languages or in none. In OT, universals hold when a structure is allowed to surface under any ranking or is ruled out under any ranking. An example of the first type is the purported phonological universal that all languages have open CV syllables. Prince and Bruce Tesar (2004, ch. 6) show that this must be the case as long as (a) no language lacks CV sequences underlyingly (see next section on this), (b) there are constraints banning onsetless and closed syllables, and (c) there are no constraints that ban onsets or require that syllables be closed. A simple example of the second type is that no language has only nasalized vowels; the presence of nasal vowels implies oral vowels. One way to derive this is by ranking the constraint against nasal vowels universally over the constraint against oral vowels (McCarthy and Prince 1995); this still predicts a grammar that has no vowels, however, and so another approach would be to assume that Con has a constraint against nasal vowels but not one that bans all oral vowels (Gouskova 2003).

21.3 OT, the lexicon, and the input

This section deals with the status of the lexicon and the input in OT, which is often a source of confusion for newcomers to the theory. As shown in §21.2.3.2, OT has no constraints that apply only to the input. Markedness constraints apply to outputs, and faithfulness constraints compare inputs and outputs. One of the motivations for this is to address the duplication problem: it is often the case that the same constraint apparently applies to derived and to underived (≈ underlying) sequences. We saw this in Tonkawa (recall (1)): there are no tautomorphemic geminates, and vowel deletion is not allowed to create new ones. An OT account explains both observations by assuming that NoGem rules out geminates at the surface level, regardless of their source. We know why Tonkawa phonology cannot create new geminates, but how does it rule out underived ones without ruling them out from the input? The answer in an OT account is that, even if hypothetical geminate inputs existed, they would not map faithfully in Tonkawa. Positing hypothetical inputs with geminates reflects an assumption known as *Richness of the Base* (Prince and Smolensky 2004): the input to a language's grammar is not subject to language-specific restrictions and may contain structures not found on the surface.

To understand Richness of the Base, it may help to distinguish between lexical entries and inputs to the grammar. OT is actually not tied to specific claims about the contents of the phonological lexicon. Much work in OT tacitly shares the SPE assumption that each morpheme has a unique underlying form that specifies its idiosyncratic features. Whether this is valid or not, lexical entries are not the same as the inputs that an OT grammar must be able to handle. A grammar describes (among other things) the speaker's knowledge of what surface forms are legal in the language, and an OT grammar derives legitimate surface forms by filtering out all illegitimate inputs. To explain why Tonkawa lacks underived geminates, we must therefore show that even if they were submitted to the grammar for evaluation, they would not map faithfully:

(26) Tonkawa and Richness of the Base: hypothetical geminate inputs map unfaithfully

/piccena-/	NoGem	Ident-Length
pi.cen∼pic.cen	W	L

Importantly, this is not a claim about the underlying representations of words in Tonkawa. The actual lexicon of Tonkawa need not have morphemes with underlying geminates, but the grammar can handle inputs with geminates nonetheless. The duplication problem is addressed here by attributing the ill-formedness of all geminates, regardless of their source, to the same constraint: NoGem.

One last point about Richness of the Base is that it is often impossible to know exactly what the hypothetical inputs map to. In the case of Tonkawa, we have a clue that (26) is on the right track, since morphologically derived geminates shorten to singleton consonants. The lack of certainty is not a grave concern, however: the analytical goal is to show that unattested structures are in fact ruled out by the analysis, and if there isn't any evidence as to their fate, the analysis simply leaves it underdetermined.

21.4 Acquisition and learnability

The advent of OT revolutionized the study of phonological acquisition because OT can directly address Jakobson's (1941) observation that child speech is less marked than adult speech. OT's constraints capture this observation directly, assuming that children start out with an initial ranking in which markedness constraints dominate faithfulness constraints (see Gnanadesikan 2004 and other contributions to Kager et al. 2004). The initial ranking idea evokes the theory of Natural Phonology (Stampe 1973), whereby children start out with universal natural rules and learn to suppress some of them. Unlike SPE, though, Natural Phonology never became influential as a theory of rules, and SPE itself had little to say about acquisition or learnability since it had no mechanism for rule learning.

Research in OT has been accompanied by parallel work on learnability almost from the very beginning (Tesar and Smolensky 2000, Prince and Tesar 2004, Hayes 1999, Boersma and Hayes 2001). To learn an OT grammar is to arrive at a constraint ranking that produces the mappings and surface forms of the target grammar without overgenerating (Prince and Tesar 2004, Hayes 2004). In a realistic setting, this would require working out not only the ranking but also underlying representations (Merchant and Tesar to appear), structural ambiguity (Tesar 1998), and other problems that must be addressed in any theory of learnability in phonology. Because the problem of learnability is so complex, research on learnability must proceed incrementally by addressing these questions one at a time, and it is still an area of ongoing work.

21.5 Variation

Variation has been the subject of keen interest in recent work in phonological theory, both in OT and in related theories such as Harmonic Grammar. This is unsurprising, since OT characterizes in an explicit way what it means for two grammars to differ. If variation is seen as the coexistence of two grammars (or subgrammars) within an individual or a community, then it naturally suggests an intuitive approach to variation: variation exists when conflicting constraints are not conclusively ranked. This can be implemented formally with only slight modifications to the basic architecture outlined in §21.2. In this section, I describe just two such approaches; the reader is invited to consult Coetzee and Pater (to appear) for an excellent recent overview of others (both in OT and other frameworks).

An influential approach to variation in OT is Partially Ordered Grammars (Kiparsky 1993, Anttila 2002). This theory revises the notion of the language-specific constraint hierarchy $\mathcal{H}$ by under-determining the rankings of crucially conflicting constraints. At the point of selecting the optimum, a specific ranking of the hierarchy must be chosen, but it is chosen at random from several alternatives. To see how this works, consider again Tonkawa. Recall that in Tonkawa compound vowel deletion may delete word-final vowels even if this brings two identical consonants together. According to Hoijer (Hoijer 1946, 1949), the consonants shorten to a single consonant, but it appears that there is some optionality to the rule. The relevant facts from (2) are repeated below.

(27) Tonkawa (de)gemination

/taʔane-nisʔo:yta-/	taʔan(n)osʔo:ta-	'to stretch (e.g., a rope)'
/yakona-nacaka-/	yakon(n)acaka-	'to kill (him) with a blow of fist'
/yakexe-xakana-/	yakex(x)akana-	'to push (it) down hard'

Importantly, this vowel deletion is obligatory, but the choice between consonant deletion and a double consonant is optional. In Partially Ordered Grammars, the constraints against consonant deletion and geminates would be tied even though they conflict with each other. At the moment of utterance, the speaker has to choose between the two rankings. The hierarchy is as follows. The constraint requiring vowel deletion at the end of the first member of the compound is FINAL-C "a prosodic word ends in a consonant," and it is categorically ranked on top. The constraints against geminates (NOGEM) and deletion (MAX-C) are tied in the next stratum, so sometimes, NOGEM will be violated (*yakex-xakana-*), and, other times, MAX-C will be violated (*yakex-akana-*). Since word-medial vowel deletion is obligatory and always blocked by NOGEM, the rest of the rankings must be fixed as shown in (30):

(28) Tonkawa optional degemination: geminated variant

	/yakexe-xakana-/	FINAL-C	MAX-C	NOGEM	MAX-V
a.☞	yakex-xakana-			*	*
b.	yakex-akana-		*!		*
c.	yakexe-xakana-	*!			

(29) Tonkawa optional degemination: degeminated variant

	/yakexe-xakana-/	FINAL-C	NOGEM	MAX-C	MAX-V
a.	yakex-xakana-		*!		*
b.☞	yakex-akana-			*	*
c.	yakexe-xakana-	*!			

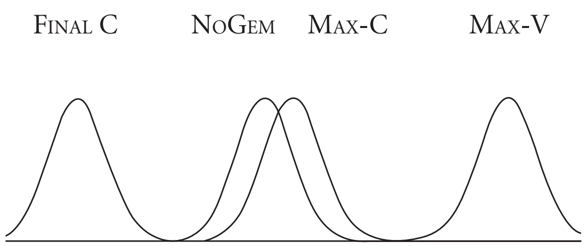
(30) Tonkawa hierarchy, in Partially Ordered Grammars

FINAL-C≫{NOGEM, MAX-C}≫STRESSED=HEAVY≫MAX-V

Partially Ordered Grammars theory adds another dimension to capturing variation. Since the choice between rankings is random, it is assumed that the frequency of variants will depend on the number of rankings of the stratified hierarchy that produce those variants (see §21.2.3.3). In this analysis of Tonkawa, consonant deletion is predicted to occur half of the time, and geminate outputs should surface the other half of the time. In this way, Partially Ordered Grammars tries to account not only for the existence of variants but also for the frequency of their distribution.

Boersma and Hayes (2001) remark that Partially Ordered Grammars can only generate variant frequencies predicted by the constraint set, which is descriptively too weak. In reality, variant frequency may depend on factors other than the grammar, and it can also be strongly skewed toward one of the variants in a way that cannot be captured with a well-motivated constraint set. They add power to their model by redefining the hierarchy as probabilistic. Any given constraint has a probability range over which it is likely to be ranked. The greater the overlap between two constraints' ranges, the higher the likelihood of ranking reversal at utterance time. The hierarchy for Tonkawa would then look something like (31), and the degree of overlap between NoGem and Max-C could be varied to match the frequency of variants should it not be 50/50%.

(31) Tonkawa hierarchy, in Stochastic OT



Boersma and Hayes also propose a learning algorithm that, they argue, is not only capable of learning the target grammar but is also robust in the face of variation and can even reproduce frequencies of variants in the target grammar. With the development of approaches like this, phonological theory can now begin to broaden its empirical scope and address variation as an aspect of phonological competence.

21.6 Conclusion

Optimality Theory has revolutionized phonological theory more than any development since the SPE. It allowed phonologists to tackle problems such as conspiracies, typological differences and universals, phonological acquisition, learnability, and variation, all by introducing the profound claim that grammars consist of violable universal constraints. Still, there are many questions that have not been answered to everyone's satisfaction. Areas of ongoing work include phonological opacity, representations, the too-many-solutions problem, and the right approaches to exceptionality, lexical stratification, and issues at the interface of phonology and phonetics and syntax. At the same time, the theory is increasingly being tested using experimental and modeling methodologies from cognitive science (see the contributions to Coetzee et al. to appear). Thus, fifteen years after its arrival, it is still a vibrant theory with many directions for development.

This page intentionally left blank

CHAPTER 22

OPTIMIZATION PRINCIPLES IN THE TYPOLOGY OF NUMBER AND ARTICLES

HENRIËTTE DE SWART
JOOST ZWARTS

22.1 Optimality Theory as a tool for typological investigations

There are about 6,000 languages spoken worldwide, most of them not mutually intelligible, and many of them displaying rich dialectal variation. Language is a key component of human cognition and highly characteristic of our interactive behavior and it is probably the only general feature that allows such a large amount of diversity across social communities and cultural groups. Why would that be and how can we get a grip on this rich inventory of linguistic variation? These questions have fascinated grammarians, linguists, philosophers of language, psychologists, and cognitive scientists of all times and places. The 20th century saw tremendous progress in our way of thinking about language and linguistic diversity, mostly because of the advent of more formal models which allowed linguists to focus on linguistic competence rather than language performance. With the creation of large corpora and electronically searchable databases, we witness a return to rich empirical data, fine-tuned observations, and statistical generalizations over actual language use. These findings show conflicting tendencies, such as convergence vs. divergence in morpho-syntactic patterns, economy and simplification vs. doubling and complex agreement phenomena. The patterns often reflect functional considerations that are well known from the typological literature, but they are tendencies, not absolute rules that are always applicable. As such, they haven't been fully integrated within more formal views on language and linguistic diversity that focus on universal grammar. We need a framework that potentially captures typological and dialectal variation, and which is grounded in descriptive adequacy as well as theoretical claims concerning patterns of linguistic form and meaning. In this chapter, we exploit the framework of Optimality Theory (OT), for it lends itself very well as a tool for typological research, and that is the aspect of the theory we focus on here.

Optimality Theory (OT) is a fairly recent development within linguistics. It originates in work on phonology (Prince and Smolensky 1997), was further developed as a framework for syntax (Barbosa et al. 1998; Legendre et al. 2001), and eventually found its way into semantics and pragmatics (de Hoop and de Swart 2000; Hendriks and de Hoop 2001; Blutner 2000; Blutner and Zeevat 2003, etc.). As the name indicates, OT is based on principles of optimization. When we speak, there are in principle infinitely many ways in which we can package our message. We pick the one that best expresses the information that we want to convey to the hearer. But in other circumstances, addressing a different person, speaking another language or another variant of the language, we would have framed our message in a different way. This intuition underlies the idea that there is no perfect form, no perfect (universal) grammar, no perfect language per se, but that we operate with locally optimal patterns of linguistic structure and language use. Variation across languages arises from differences in the "weights" assigned to certain symbolic rules (phonological, morphological, syntactic, semantic ones). These rules represent different, possibly conflicting tendencies in language. In particular, OT exploits the opposition between a drive for economy (favoring "simpler" expressions over more complex ones) and a drive for reflecting differences in meaning in the form. All rules are soft, and can be violated if necessary to satisfy rules that are considered more important in the language, and are ranked higher in the grammar. Languages vary in the balance they strike between economy and meaningful formal distinctions, which is reflected in a range of linguistic patterns.

In this chapter, we will illustrate the essence of such an approach with a typology of plural morphology and article use. Our goal is not to give a complete overview of how these grammatical categories are expressed cross-linguistically but rather to demonstrate how certain patterns of variation can be understood from the theoretical perspective of Optimality Theory. The core question we address is: what is the best way for the speaker to convey in his or her language information concerning singular/plural distinctions, definiteness, and discourse reference in the nominal form? OT frames this as a relation between input and output, and spells out an optimization process over candidates that are potential forms to convey a certain message. If the speaker wants to convey a certain meaning (input), we determine what is the optimal form (output) he or she should use to do so under the language-specific grammar, which is conceived of as a ranked set of constraints. The meaning carried by the form has to be recoverable, that is, the hearer must be able to construe a meaningful representation based on the form chosen by the speaker. Successful communication arises when this meaning representation corresponds to the message the speaker intended to convey. Our approach is both deductive and partial: starting from a few theoretical premises, we build a simple model and then test how it accounts for part of the cross-linguistic data (section 22.2). Extensions enlarge the empirical coverage and open up a wider perspective (sections 22.3 and 22.4). These sections function as a guide for possible future research.

22.2 Articles and plurality across languages: an OT typology

In this section, we demonstrate how a typology is set up in Optimality Theory. We define a single economy constraint, and a range of constraints driving the expression of distinctions in number, definiteness, and discourse reference. Reranking of constraints leads to a range of possible languages, and provides a basic account of how language variation arises.

22.2.1 A general markedness constraint: *FunctN

Plural morphology and definite/indefinite articles convey grammatical or "functional" information, as opposed to the lexical information contributed by nouns. In this chapter we ignore proper names and pronominal structures, and focus on expressions with a lexical core, conveyed by a noun (N). East-Asian languages such as Mandarin Chinese, Japanese, Korean make extensive use of classifier and quantifier constructions, but do not systematically establish singular/plural distinctions on nouns, nor do they use definite/indefinite articles like English *a* and *the*.

Accordingly, a sentence like (1) (from Krifka 1995) can be paraphrased by different English sentences:

(1) Wò kànjiàn xióng le [Mandarin Chinese]
I see bear ASP
'I see a bear/some bears/the bear(s).'

As illustrated by (1), Mandarin Chinese allows "bare" nominals, i.e., nominals without articles, without plural morphology, without any sort of grammatical marking on the noun. Such bare nominals are preferred under a general economy constraint barring (morphologically and syntactically) complex nominals involving grammatical or functional adornment of N. We label this constraint *FunctN:

- *FunctN: Avoid functional structure in the nominal domain.

The set of functional expressions in the nominal domain is quite rich. Besides plural morphology and definite/indefinite articles, we find demonstratives (*this/that*), classifiers (familiar from East-Asian languages such as Korean, Japanese, Chinese), case marking, etc. We will not formalize the general notion of "functional structure" in a specific linguistic framework in this chapter, but assume that all parts of a nominal constituent accompanying the lexical core are involved. *FunctN is called a *markedness* or *economy* constraint because it favors nominals with the least complex nominal structure. The intuition underlying markedness constraints is that certain forms are preferred to other forms because they are simpler, or shorter, or occur in more languages. In line with the idea that markedness constraints bar complexities, they are often formulated as "avoid" constraints, and we see this with "Avoid functional structure in the nominal domain". *FunctN is a gradable constraint, as each grammatical marker (plural morphology, article, etc.) adds to the complexity of the form.

The concept of markedness is not always easy to define, and different notions of markedness are found in the literature (cf. Haspelmath 2006 for discussion). However, markedness or economy is broadly conceived as a factor shaping language by many different theoretical frameworks, ranging from functionally motivated approaches to minimalist syntax. In OT, markedness defines a set of output oriented constraints: no matter what message the speaker intends to convey, he or she will prefer a simpler over a more complex form to do so. It is out of the scope of this chapter to address the full inventory of grammatical markers in the nominal domain, so we focus on plural morphology and definite/indefinite articles.

*FunctN is a powerful markedness constraint, but economy is not the only force driving natural language. The tendency toward simplification is often in conflict with devices that elaborate form and multiply markings. Such opposing tendencies are difficult to manage in models of linguistic theory that are built on hard rules that always need to be satisfied. One of the innovative ingredients of an optimization approach to grammar is the assumption that grammatical rules are not absolute but can be violated if necessary. The Mandarin Chinese example in (1) above indicates that *FunctN occupies a high position in the ranked set of constraints that constitutes the grammar of the language. In the remainder of this section, we consider cases in which the economy constraint *FunctN is overruled by constraints that favor the encoding of distinctions in meaning in the functional structure of the nominal. If other constraints outrank *FunctN, we obtain a richer set of possible nominal structures in the language, with concomitant meanings. In sections 22.2.2 through 22.2.4, we outline a set of so-called faithfulness constraints which enforce the expression of number distinctions and definite/indefinite discourse reference in the form of the nominal.

22.2.2 Plural morphology driven by Fpl

Unlike Mandarin Chinese, many languages mark singular/plural distinctions on the noun by means of special number morphology. (2) provides an example from Polish (D. Klimek, p.c.).

(2) a. Piłka toczy się.
ball roll-3sg refl.
'A/The ball is rolling.'

b. Piłki toczą się.
balls roll-3pl refl.
'Balls/The balls are rolling.'

Polish does not use articles to distinguish definite from indefinite noun phrases but it does establish a singular/plural distinction in the morphology of the noun. From the typological literature, the generalization emerges that plural will be marked first, if there is a number distinction at all (Greenberg 1966*b*; Corbett 2000).[1] The unmarked form is then used for singular reference (and possibly number neutrality, cf. Farkas and de Swart 2003). The form marked with a plural morpheme refers to a group of individuals (Farkas and de Swart 2003, 2007). We formulate this as a constraint between the input (the information the speaker wants to convey) and the output (the shape of the nominal constituent):

- Fpl: reference to a group of individuals must be reflected in a special plural form of the nominal.

The special plural form of the nominal usually involves plural morphology on the noun, but other realizations of plural are possible (cf. Rijkhoff 2002; Corbett 2000). In OT terms, Fpl is a *faithfulness* constraint. Faithfulness constraints link a particular input (in this case reference to a group of individuals) to a certain output (in this case, a special plural form of the nominal). The special plural marking signals to the hearer that the speaker is talking about a group of individuals.

[1] We will come back to the asymmetry between singular and plural marking in section 22.3 below.

Tableau 22.1. No plural marking in Mandarin Chinese

Meaning ∃x_pl Bear(x) & See(I, x)	Form	*FunctN	Fpl
☞	Wò kànjiàn xióng le I see bear ASP		*
	Wò kànjiàn xióng_pl le I see bear ASP	*	

Tableau 22.2. Plural marking in Polish

Meaning ∃x_pl Ball(x) & Roll(x)	Form	Fpl	*FunctN
	Piłka toczy się. Ball roll-3sg refl.	*	
☞	Piłki toczą się. Balls roll-3pl refl.		*

Faithfulness constraints often favor more complex output candidates. Given that it promotes the use of special grammatical structure in the nominal domain, the faithfulness constraint Fpl is in conflict with the markedness constraint *FunctN, which bars such functional markers, and favors simpler forms. Given that OT constraints are violable, we can resolve the conflict by differentiating the weight of opposing constraints. In the interaction of two constraints (call them C_1 and C_2), there are two possible rankings: either C_1 is more important, and outranks C_2 (written as $C_1 \gg C_2$), or C_2 outranks C_1 (written as $C_2 \gg C_1$). A grammar that ranks *FunctN above Fpl (*FunctN ≫ Fpl) captures East-Asian languages such as Mandarin Chinese, which produce the same form for singular and plural (example 1).[2] The description of Polish involves a grammar that ranks Fpl above *FunctN (Fpl ≫ *FunctN). Such languages produce a special form for plural nominals (example 2). The optimization process is spelled out in the OT Tableaux 22.1 and 22.2.

As the input to the optimization process, the first column of the tableau gives a representation of the intended meaning in terms of a first-order logical formula, enriched with information concerning group-level reference of the object. There is a long, possibly infinite set of forms that could convey indefinite reference to a group of balls, a selection of which are listed as candidate forms in the second column. In the Tableaux 22.1 and 22.2, we focus on the two candidates that differ in the presence vs. absence of special plural morphology on the noun. The remaining columns list the constraints in order of strength. Constraints more to the left are stronger than constraints more to the right. The ideal candidate does not violate

[2] Mandarin Chinese has an optional plural classifier, but this expression is not grammaticized as a general plural marker, obligatorily used for plural reference; cf. Cheng and Sybesma (1999) for discussion.

any constraints, but in Tableaux 22.1 and 22.2, each candidate violates a different constraint (marked with an asterisk *). The conflict between economy of form and faithfulness to number distinctions is resolved by the ranking of the constraints. Under an optimization approach, violations of a lower-ranked constraint are tolerated if such a pattern allows satisfaction of a higher-ranked constraint. The optimal candidate under a particular constraint ranking is indicated by the pointing hand (☞). Mandarin Chinese and Polish crucially differ in the ranking of *FUNCTN and FPL. As a result, the optimal form in Mandarin Chinese is a nominal not marked for plural reference (Tableau 22.1), whereas in Polish, the optimal candidate reflects group reference in a special plural form of the nominal (Tableau 22.2). Note that there are more violations of *FUNCTN than indicated in Tableau 22.3 (in particular for case marking), but we only indicate those relevant to the distinction in number.

The markedness constraint *FUNCTN reflects speaker economy (it is "easier" to produce an unmarked form), whereas the faithfulness constraint FPL reflects the hearer's perspective (it is "easier" to understand that group reference is intended when a specially marked form conveys this information). The contrast between the two languages shows that typological variation is the result of conflicting tendencies in natural language, which allow more than one solution. Along similar lines, the speaker can convey information concerning referential status of the individuals talked about by means of definite and indefinite articles. In the next section, we will be concerned with discourse referential status, and uniqueness/familiarity of the referent. The relevant distinctions made in natural language are captured by the two faithfulness constraints FDEF and FDR introduced in section 22.2.3.

22.2.3 Article use driven by FDEF and FDR

Many languages have a definite article conveying uniqueness ("the queen of the Netherlands" refers to the one and only queen of the Netherlands), maximality of groups of individuals ("the stars" refers to all the stars) or familiarity (in "I saw a dog in the park. The dog wagged its tail", *the dog* refers to the dog I just saw). We abstract away from the question whether the definite article involves a condition on uniqueness/maximality (as proposed by Strawson 1950, Hawkins 1991), or should be framed in terms of the discourse old/new distinction (Heim 1982), and use the term "discourse uniqueness" to generalize over both uses.[3]

- FDEF: Reference to discourse unique individuals (unique/maximal or familiar ones) requires the use of an expression of definiteness.

In this chapter, we will limit ourselves to languages using definite articles, which constitute the typical expression of definiteness, but the constraint FDEF does not

[3] The notion of discourse uniqueness is inspired by Farkas (2002), who develops a formal implementation of such a unified view in Discourse Representation theory.

Tableau 22.3. Bare definite plurals in Hindi

Meaning ∃!x_pl Child(x) & Play(x)	Form	FPL	*FUNCTN	FDEF
	bacca khel rahaa hai child play PROG PRES	*		*
☞	bacce khel rahe haiN children play PROG PRES		*	*
	DEF-bacce khel rahe haiN the children play PROG PRES		**	

Tableau 22.4. Definite article in Hebrew

Meaning ∃!x Tiger(x) & Disappear(x)	Form	FDEF	FPL	*FUNCTN
	namer ne'elam tiger disappeared	*		
☞	ha-namer ne'elam the-tiger disappeared			*

exclude other means of conveying definiteness (compare analyses of case-marking in Turkish and Persian in terms of specificity by Enç 1991 and Karimi 2003). In the grammar of Mandarin Chinese, Polish, or Hindi, *FUNCTN outranks the faithfulness constraint FDEF, so we do not find definite articles. In such languages, bare nominals display ambiguities between definite and indefinite readings, as illustrated by the Hindi bare plural in (3) (from Dayal 1999):[4]

(3) bacce khel rahe haiN [Hindi]
children play PROG PRES
'The children/some children are playing.'

With the ranking FDEF ≫ *FUNCTN, we obtain a system in which definite articles alternate with bare nominals, as we see in Bulgarian and Hebrew. The Hebrew example in (4a) is from Doron (2003), who ascribes it a regular definite or a generic interpretation. It contrasts with (4b) in which the bare singular is ascribed an indefinite (non-generic) interpretation:

(4) a. ha-namer ne'elam me ezor-enu. [Hebrew]
the-tiger disappeared from area-our
'The tiger disappeared from our area.'

b. namer ne'elam me ezor-enu.
tiger disappeared from area-our
'A tiger disappeared from our area.'

[4] The situation of Hindi bare singulars is more complex, for the choice between a definite and an indefinite interpretation is affected by other syntactic and semantic particularities; cf. Dayal (1999; 2004) for discussion.

Tableaux 22.3 and 22.4 spell out the optimization patterns in Hindi and Hebrew respectively.

Hindi establishes a distinction between singular and plural nominals, just like Polish, so it has the ranking FPL ≫ *FUNCTN (cf. Tableau 22.2). There is no definite article, so we posit a grammar for Hindi in which *FUNCTN is ranked above FDEF. Under this ranking, economy is more important than faithfulness to the discourse referential distinction between discourse unique or discourse non-unique individuals. Although the input in Tableau 22.3 is a unique maximal group of individuals (marked as ∃!x_pl), the optimal form is a bare nominal.

Just like Polish and Hindi, Hebrew establishes a singular/plural distinction between nouns, so its grammar shows the ranking FPL ≫ *FUNCTN. Unlike these other languages, Hebrew contrasts definite and bare nominals, so the faithfulness constraint FDEF is ranked above the markedness constraint *FUNCTN in Tableau 22.4. The speaker wants to convey information about a unique individual, so the input semantic representation marks the individual referred to as ∃!x. Faithfulness to uniqueness/discourse familiarity requires marking of the input meaning by means of a definite article in the form. The optimal nominal structure conveying the meaning intended by the speaker under the ranking posited in Tableau 22.4 is the definite description *ha-namer*. Note that the respective order of FPL and FDEF does not matter, for the two faithfulness constraints do not interact. In the tableau, this is reflected by the dotted line between the two columns.

The introduction of faithfulness constraints concerning the referential status of the discourse referent should be viewed in the broader perspective of how discourse referents are set up by linguistic expressions. It is generally acknowledged that articles, quantifiers, and numerals (broadly construed as a category of determiners) are used to ground the individuals described by the noun in the conversational context. Nominals in regular subject/object/indirect object position are assumed to have full discourse referential status (cf. Higginbotham 1985; Kamp and van Eijck 1996; Chierchia 1998). In the absence of further information concerning definiteness, number or quantificational status, such nominals take an indefinite article. Bare (singular) nominals are blocked in regular argument position in languages like English, Dutch, German, etc. (5a). However, nominals in incorporation constructions, predicative contexts, and a range of other constructions are non-discourse referential, and we observe that bare nominals are allowed in such environments, albeit in a restricted way (5b).[5]

(5) a. Susan ate *apple/an apple.
    b. Susan is head of the department of linguistics.

[5] The reader is referred to Farkas and de Swart (2003) and literature cited there for claims concerning the lack of discourse referentiality for incorporated nominals, to de Swart et al. (2005, 2007) for similar claims concerning predicative nominals, and to de Swart and Zwarts (2009) for a more general discussion of non-referentiality.

The correlation between determiners and discourse referential status is captured by the constraint Fdr:

- Fdr (preliminary version): The presence of a discourse referent in the semantics corresponds with the presence of a determiner in the nominal form.

Note that Fdr is independent of the constraint Fdef, in that it does not convey information concerning discourse uniqueness. The contrast between definite and indefinite articles in English makes this a bit hard to see, but the Salish languages provide a nice illustration. Matthewson (1998) discusses the following examples from St'át'imcets:

(6) a. tecwp-mín-lhkan ti púkw-a lhkúnsa. [St'át'imcets]
buy.appl-1sg.sub det book-det today
'I bought a/the book today.'

b. Léxlex I smelhmúlhats-a.
Intelligent det.pl woman.pl-det
'Women/the women are intelligent.'

According to Matthewson, the St'át'imcets (circumfixed) determiners do not encode either definiteness or specificity but merely assert existence. She models this notion in Discourse Representation Theory (Kamp and Reyle 1993). For our purposes, "assertion of existence" can be identified with the introduction of a discourse referent. The ranking of Fdr above *FunctN forces the introduction of a determiner when the nominal is in regular argument position. Further evidence that Fdr is at stake in the grammar of St'át'imcets comes from the observation that the determiner does not appear in predicative contexts, compare (7a) and (b).

(7) a. kúkwpi7 kw s-Rose.
Chief det nom-Rose
'Rose is a chief.'

b. *ti kúkwpi7-a kw s-Rose.
det chief-det det nom-Rose

Under the assumption that predicate nominals do not have discourse referential force (cf. de Swart et al. 2005, 2007), the semantic input in (7) does not contain a discourse referent, so the use of the determiner is not enforced. Given that the language does not establish a definite/indefinite contrast, we posit a low position for Fdef in the grammar. The ranking {Fpl, Fdr} ≫ *FunctN ≫ Fdef derives the obligatory use of a definite-neutral article for all nominals in regular argument position (Tableau 22.5). Predicate nominals occur bare, as illustrated in Tableau 22.6.

The bare nominal violates the highly ranked constraint Fdr, so under the ranking Fdr ≫ *FunctN in St'át'imcets, the nominal in direct object position is obligatorily

Tableau 22.5. Discourse reference in St'át'imcets

Meaning ∃!x Book(x) & Buy(I, x)	Form	FDR	FPL	*FUNCTN	FDEF
	tecwp-mín-lhkan púkw buy.APPL-1 SG.SUB book	*			*
☞	tecwp-mín-lhkan ti púkw-a buy.APPL-1 SG.SUB DET book-DET			*	*

Tableau 22.6. Bare nominals in predicative contexts in St'át'imcets

Meaning Chief(rose)	Form	FDR	FPL	*FUNCTN	FDEF
☞	kúkwpi7 kw s-Rose chief DET nom-Rose				
	ti kúkwpi7-a kw s-Rose DET chief-DET DET nom-Rose			*	

marked with a determiner indicating discourse referential force. As Tableau 22.5 illustrates, non-definite and definite meanings are expressed by the same determiner form. In predicative contexts, no discourse referent is present in the semantic input, so in Tableau 22.6, the bare nominal does not violate FDR. Even though the markedness constraint *FUNCTN is ranked fairly low in St'át'imcets, it comes into play in this configuration, and favors the use of a bare nominal over a nominal marked with a determiner (cf. 7a versus 7b). In OT, the phenomenon according to which a low-ranked constraint can be influential in situations in which higher-ranked constraints do not decide between candidates is known as "the emergence of the unmarked". It implies that unmarkedness is pervasive, even in languages in which high-ranked faithfulness constraints frequently block the simplest form.

St'át'imcets obligatorily marks both singular and plural nominals with a determiner indicating discourse referential status in environments such as (6). Languages like English display a contrast between singulars and plurals. Singular bare nominals do not occur in argument position (5a repeated as 8a), but bare plurals are perfectly felicitous there (8b):

(8) a. Susan ate *apple/an apple.
  b. Susan ate apples.

According to Farkas and de Swart (2003), plural reference can only be predicated of full-fledged discourse referents. Plural morphology then presupposes discourse referential status of the nominal expression. In languages like English, this allows plural morphology to introduce discourse referents through the backdoor of presupposition accommodation. When the speaker uses a presupposition triggering expression, the hearer can accommodate this presupposition in certain contexts, as illustrated by the example in (9):

(9) (teacher, upon entering the classroom late): Sorry I am late. My bike had a flat tire.

The students may be quite unaware of the fact that the teacher bikes to class, but will not dispute the existential presupposition introduced by *my bike*. The statement that the teacher owns a bike is tacitly added to the background knowledge shared by teacher and student. If this process of presupposition accommodation is operative with plural morphology, it will allow the hearer to tacitly add a discourse referent to the list of individuals talked about during the conversation, even in the absence of a determiner (cf. 8b).

In the following, we will use the term “strong” plural morphology to describe languages like English, in which bare plurals can be used with full discourse referential status, and “weak” plural morphology to describe languages in which plural inflection has to agree with a plural determiner (St’át’imcets, 6b). The correlation between strong plural morphology, determiners and discourse referential status is captured by the revised version of the constraint FDR:

- FDR (final): The presence of a discourse referent in the semantics corresponds with an expression that carries discourse referential force (“strong” plural morphology, an article, or another determiner).

A high ranking of FDR in the grammar implies that nominals in regular argument position are never fully bare (i.e., consisting solely of a lexical core).

Languages that rank FPL, FDEF ànd FDR above *FUNCTN display a full contrast between singular and plural, definite and indefinite nominals. In the singular, we find a definite and an indefinite article (10a, 11a). Depending on whether the language has strong plural morphology (e.g., English) or weak plural morphology (e.g., French, cf. Delfitto and Schroten 1991), we find a contrast between definite and bare plurals (10b) or a contrast between definite and indefinite plurals (11b):

(10) a. I bought a book/the book today.
b. I bought books/the books today.

(11) a. J’ai acheté un/le livre aujourd’hui. [French]
I-have bought INDEF.SG/DEF.SG book today
‘I have bought a/the book today.’

b. J’ai acheté des/les livres aujourd’hui.
I-have bought INDEF.PL/DEF.PL books today
‘I have bought books/the books today.’

The constraint ranking for the two languages is the same, namely {FPL, FDEF, FDR} $\gg$ *FUNCTN, where the mutual ranking of the constraints conjoined by bracketing is irrelevant. As the contrast between Tableaux 22.7 and 22.8 shows, the weak plural morphology we find in French implies a lack of satisfaction of FDR by bare plurals. Accordingly, an indefinite plural article must be inserted here.

Tableau 22.7. Indefinite singulars and indefinite plurals in French

Meaning ∃x book(x) & Buy(I, x)	Form	FDEF	FDR	FPL	*FUNCTN
	J'ai acheté livre		*		
☞	J'ai acheté un livre				*
Meaning ∃x_pl book(x) & Buy(I, x)	Form				
	J'ai acheté livres		*		*
☞	J'ai acheté des livres				**

Tableau 22.8. Indefinite singulars and bare plurals in English

Meaning ∃x book(x) & Buy(I, x)	Form	FDEF	FDR	FPL	*FUNCTN
	I bought book		*		
☞	I bought a book				*
Meaning ∃x_pl book(x) & Buy(I, x)	Form				
☞	I bought books				*
	I bought indef_pl books				**

The strong plural morphology we find in English ensures that FDR is satisfied by the plural nominal. Insertion of a plural indefinite article is thus redundant, and penalized by the markedness constraint *FUNCTN. Since there is no plural indefinite article in English, we have represented this option here with the abstract marker indef_pl.

The use of a bare, rather than a full indefinite plural in English illustrates the relevance of economy, even in languages in which *Funct is ranked fairly low. Just like in St'át'imcets bare predicative nominals, we see the emergence of the unmarked at work here.

22.2.4 Summing up: An OT typology of plural morphology and article use

The introduction of a core markedness constraint *FUNCTN and the faithfulness constraints FPL, FDEF, and FDR complete the set of constraints we discuss for our typology in this section. What emerges is a typology of languages based on all the possible rankings of these four constraints. Table 22.1 sums up the rankings exemplified so far.

Table 22.1. OT typology of plural morphology and article use

Ranking	Characteristics	Example language
*FunctN ≫ {Fpl, Fdef, Fdr}	no number morphology, no articles	Mandarin Chinese
Fpl ≫ *FunctN ≫ {Fdef, Fdr}	sg/pl distinction, no articles	Hindi, Polish
{Fpl, Fdef} ≫ *FunctN ≫ Fdr	sg/pl distinction; definite/bare contrast in sg and pl	Hebrew, Bulgarian
{Fpl, Fdr} ≫ *FunctN ≫ Fdef	no def/indef; no bare nominals (weak plural)	St'át'imcets
{Fpl, Fdr, Fdef } ≫ *FunctN	def/indef in sg; definite/bare in plural (strong plural)	English, German
{Fpl, Fdr, Fdef} ≫ *FunctN	def/indef in sg and plural; no bare nominals (weak plural)	French

From this table, Mandarin Chinese emerges as the most economical language, but that is of course influenced by the fact that we did not work out in this section faithfulness constraints licensing classifiers and other grammatical markers that occur in East Asian languages. Once we develop such constraints, they will be posited above *FunctN, thus permitting more complex nominals, even in languages that do not use articles, and do not use number morphology on the noun. The rankings in the first column show that the faithfulness constraints Fpl, Fdef, and Fdr do not interact with each other. In principle, a factorial typology based on four constraints would lead to 24 possible languages, but the lack of interaction between the three faithfulness constraints reduces this number to eight equivalence classes, five of which are distinguished in Table 22.1. Furthermore, the distinction between strong and weak plural morphology creates two options for all languages in which both Fdr and Fpl are ranked above *FunctN. This accounts for the contrast between English and French, in the last two rows of Table 22.1. So far, we haven't found a language that behaves like St'átimcets, but has a strong plural, although the case of Sinhala (section 22.3.2 below) comes very close. What is most striking about Table 22.1 is that we are missing languages in which referential faithfulness constraints are rising above *FunctN, while Fpl remains below that markedness constraint. We are not aware of languages that establish definite/indefinite and/or discourse referential distinctions in the absence of a singular/plural distinction. This might suggest that somehow number distinctions are more "basic", and are established prior to discourse referential distinctions, but further empirical research is needed to confirm this.

As we already indicated above, number and article systems across languages are more complex than what has been discussed so far. We provide more in-depth discussion of singular/plural marking in section 22.3, and an elaboration of the article system in section 22.4.

22.3 Markedness reversal in OT

In this section, we discuss two instances of markedness reversal, and a possible strategy to deal with such patterns in Optimality Theory. The first case involves unexpected patterns of number marking (section 22.3.1), the second a Sinhala article marking indefiniteness in the absence of a definite article (section 22.3.2).

22.3.1 Unexpected patterns of number marking

In the preceding section we have seen two types of languages with respect to number marking, defined by the two possible rankings of the markedness constraint *FunctN and the faithfulness-to-plural constraint Fpl:

(12) a. languages without number marking: *FunctN ≫ Fpl
b. languages with plural marking: Fpl ≫ *FunctN

In the first type of language, there is no opposition between singular and plural. This is the situation in Mandarin Chinese (cf. example 1, Tableau 22.1). In the second type of language, the plural is marked in opposition to a bare form, used for singular reference. This is what typically happens in English, as shown in the following two tableaux (which ignore the referential faithfulness constraints). The ending *-sg* in Tableau 22.9 represents the non-existent singular marker in English. That it does not exist follows from the absence of a highly ranked faithfulness constraint for singular.

However, in English we actually find Chinese patterns, too, nouns that do not contrast singular and plural. This happens for isolated cases like *sheep*, but also more regularly with certain classes of nouns, e.g., those for exotic people groups, animals that are hunted or fished on, and deadjectival human terms. Notice how the subjects in the following examples behave as plurals (triggering plural agreement on the verb), without being morphologically marked as such:

(13) a. The Carib were noted for their ferocity.
b. Carp breed from May to July.
c. The Chinese are subsidizing the American way of life.

This is a serious problem for the standard view on markedness in which there is a rigid asymmetry between singular and plural, with singular unmarked and plural marked. We find a general markedness pattern in English (and across languages), but the pattern has local exceptions, where the markedness is neutralized or reversed. This phenomenon was discussed in Tiersma (1982), who observed many more cases of *local markedness*, as he called it, arguing that in the domain of number it concerns “nouns whose referents naturally occur in groups or pairs”

Tableau 22.9. Singular as unmarked form in English

Meaning ∃x book(x)	Form	FPL	*FUNCTN
☞	book		
	book-sg		*

Tableau 22.10. Plural as marked form in English

Meaning ∃x_pl book(x)	Form	FPL	*FUNCTN
	book	*	
☞	books		*

(Tiersma 1982: 832). In a recent critical discussion of the notion of markedness, Haspelmath (2006) uses the phenomenon of local markedness as an argument for a frequency-based approach to such patterns. The plurals in (13) remain unmarked because they are used more frequently than their singular counterparts.

How exactly frequency influences marking patterns is still an open question, but the important thing is that Optimality Theory provides a framework in which local markedness can be modeled through the interaction of general and local constraints. The architecture of the OT system allows for "exceptions" and reversals in an elegant way. We illustrate that here for a few examples in the number domain, but we believe that the OT tool lends itself to application to a much wider range of typological phenomena involving markedness reversals.

Until now we have worked with one faithfulness constraint for plurality (FPL) and one markedness constraint for functional structure that gives us the results in Tableaux 22.9 and 22.10. When we encounter a situation where the plural gets no marking, as illustrated by *carp* in (13b), there is a simple way to accomplish this in the current setup: by introducing a higher-ranked markedness constraint that blocks plural marking. We would like to suggest that this constraint has a general part (*PL, "don't use plural marking"), but that it is restricted to applying to a particular class of nouns only, namely names for fish (*PL_{FISH}). Now, assuming that *carp* is a member of this semantically defined FISH class, we get the result in Tableau 22.11:

Tableau 22.11. Unmarked plural in English

Meaning ∃x_pl carp(x)	Form	*PL_{FISH}	FPL	*FUNCTN
☞	carp		*	
	carps	*		*

Since the constraint *PL_{FISH} is ranked over the faithfulness constraint FPL it forces the selection of the unmarked form *carp*. The *general* markedness constraint *FUNCTN is ranked below FPL, the *local* markedness constraint *PL_{FISH} is ranked above it, allowing for certain nouns to systematically escape plural marking. *PL_{FISH} can be seen as a very specific instance of the general markedness constraint *FUNCTN, which in English has moved upward in the constraint ranking, but in other languages has remained included in *FUNCTN.

Let us now consider a slightly more complicated language in which not only plural nouns can remain unmarked but singular nouns can also be marked. The singular marking is often called *singulative*. This is especially common in Eastern Africa, in both Nilo-Saharan and Afro-Asiatic languages (Dimmendaal 2000). Here are some examples from Endo (Nilotic, Nilo-Saharan, Zwarts 2007):[6]

(14) a. chumpa 'Europeans'–chümpïïn 'European'
b. taraak 'cedars'–tärääkwä 'cedar'
c. taalim 'grasshoppers'–taalimwa 'grasshopper'
d. pëël 'elephants'–pëëlyoon 'elephant'
e. mur 'rats'–muryaan 'rat'

In this pattern the plural form is unmarked, the singular form has a special singular suffix. This language does not mark all nouns in this way, but, on the other hand, these are no rare exceptions, for they involve a sizable proportion of the nouns and systematic patterns can be distinguished. Most of the people groups, trees, and insects are treated like this, but there are also isolated cases, as for example the elephant (which is unlike the other bigger mammals, which are unmarked in the singular, marked in the plural). The general idea in the literature (e.g., Dimmendaal 2000) is that these patterns are based on the fact that for these nouns the plural is more frequent than the singular. Phrased differently: Europeans, cedars, grasshoppers, elephants, and rats occur more often in groups than as individuals.

So, it seems that the same "multitude" factor that plays a role in neutralizing the plural in English is also working in Endo, but note that Endo differs from English in a crucial way. Unlike in English, the singular in Endo is not left unmarked, but marked with a singulative suffix. In fact, there are hardly any nouns in Endo that have the same form for singular and plural. We see that the same functional factors are working in English and Endo, but that they are organized in different ways, leading to typologically different systems.

Let's see how OT can deal with this, on the basis of the example *taraak* "cedars" vs. *tärääkwä* "cedar". The majority of nouns in Endo marks only plural, so we assume the FPL $\gg$ *FUNCTN order that we also see in English. With this ordering we would expect the plural of *taraak* to be marked, so we need a local markedness constraint. We use a higher-ranked markedness constraint *PL_{TREE} ("leave the plural of trees

[6] The diaeresis expresses advanced tongue root. Tone is left unmarked.

Tableau 22.12. Unmarked plural in Endo

Meaning ∃x_pl cedar(x)	Form	FSG$_{TREE}$	*PL$_{TREE}$	FPL	*FUNCTN
☞	taraak			*	
	taraaktiin		*		*

Tableau 22.13. Marked singular in Endo

Meaning ∃x cedar(x)	Form	FSG$_{TREE}$	*PL$_{TREE}$	FPL	*FUNCTN
	taraak	*			
☞	täräákwä				*

unmarked"), because that seems to be the subregularity in the number system. For the singular we now need a special faithfulness constraint that guarantees the marking of the singular for the class of trees, FSg$_{TREE}$ ("mark the singular of trees"). These two constraints go hand in hand, as two sides of the same coin, making sure that number of trees is always differentiated in Endo. This is illustrated in Tableaux 22.13 where -*tiin* is a plural suffix and -*wä* is the singular suffix.

Tableau 22.12 captures why *taraak* "cedar" does not have the plural form *taraak-tiin* when it refers to a plurality (or any of the many other possible plural forms that Endo morphology provides). What makes Endo *taraak* different from English *carp* is that Endo requires number to be differentiated. When the plural is neutralized for a particular class of nouns (through a functionally motivated markedness constraint), then the singular needs to be reflected in the output to avoid the distinction being lost, as shown in Tableau 22.13. This is why in Endo *PL$_X$ for a particular lexical class X will pair up with FSG$_X$ for that class.

Number marking in Endo, and in languages across the world, is much more complex than this, but this first sketch gives us an idea of how constraint interaction could help us understand some of the complexity.

22.3.2 An indefinite article in the absence of a definiteness marker

When a language has only one type of article (definite vs. indefinite), then, according to typological generalizations (Dryer 2007), this article is more likely to be definite than indefinite. In the OT system, this is reflected by the presence of a constraint FDEF, rather than a constraint FINDEF. Indefinite articles in languages like English are licensed by the constraint FDR, and get an indefinite semantics, because they take the complementary meaning of the definite article. This view of indefiniteness as non-definiteness is challenged by languages that have developed an indefinite article in the absence of a marker of definiteness. Sinhala instantiates this reversed markedness pattern.[7] Sinhala has a high ranking for the constraint FPL, and establishes a systematic morphological distinction between singular and plural nouns. It does not have a definite article. Plural nouns are bare, regardless of whether the noun is definite or indefinite. However, in the singular, indefiniteness is marked by the article *-ek* (for animates) or *-ak* (for inanimates). In the absence of an indefinite article, the bare singular noun is interpreted as definite. So for instance *vacanaya* or *vacane* is the bare form, meaning "the word", *vacana* is the plural form meaning "words" or "the words", and *vacanayak* is the singular indefinite form, meaning "a word". Examples in context (from Henderson 2006) are given in (15).

(15) a. waňdura kehel malə uḍə-ṭə pænn-a [Sinhala]
monkey banana flower.SG top-DAT jump-PST
'The monkey jumped on the stalk of bananas.'

b. mæssa miris karələkə wæhuw-a
fly.SG chili.pepper pod.like.thing.SG.IND.LOC land-PST
'The fly landed on a chili pepper.'

c. hatu mal narakweela
mushroom flower.PL rotten
'The mushrooms are rotten.'

d. maŋ laŋgə dehi geḍi tiye-nəwa
1SG near lime fruit.PL exist-IMPF
'I have limes.'

The indefinite marker *-ek* is identical to the stem of the numeral "one", but just like the numerals "two", "three", etc., "one" comes in two forms, namely the definite form *ekə* (inanimate) or *ekenaa* (animate), and the indefinite form *ekak* (inanimate) or *ekkenek/kenek* (Henderson 2006, with reference to Gair and Paolillo 1997: 22). Thus no reduction of the indefinite article to the numeral "one" is possible. We can account for the reversed markedness strategy adopted by Sinhala along similar lines as the reverse number systems in section 22.3.1.

Suppose Sinhala wants to mark both plural and discourse referentiality (similar to St'át'imcets), but differs from the Salish languages in blocking a general marker of definiteness. This suggests the ranking *DEF ≫ {FDR, FPL} ≫ *FUNCTN ≫ FDEF. Strong plural morphology is sufficient to satisfy the faithfulness constraint FDR with plural nominals, and the high ranking of *DEF implies lack of differentiation between definite and indefinite interpretations of the bare plural. Hijacking the numeral "one" as a marker of indefiniteness satisfies FDR in the singular domain,

[7] Thanks to Gavin Austin (p.c.) for drawing our attention to the relevance of Sinhala for our analysis. (cf. Austin (2008).

Tableau 22.14. Indefinite singular in Sinhala

Meaning ∃!x word(x)	Form	*DEF	FDR	FPL	*FUNCTN	FDEF
☞	vacane word		*			*
	vacane.DEF word.DEF	*			*	
Meaning ∃x word(x)	Form					
	vacane word		*			
☞	*vacanayak* word.INDEF				*	

but an overt definite form is blocked by the high ranking of *DEF. Accordingly, Sinhala reverts to the unmarked bare form to convey definiteness. The paradigm is illustrated in Tableau 22.14.

Summing up, we observe that OT often allows for different choices in setting up constraint systems and in different architectures of the form–meaning correspondence. However, the important thing is that OT is a formal framework in which the different factors that shape grammatical marking can be brought together as sharply formulated constraints on the mapping between meaning and form. More specifically, we have seen how functionally grounded factors in local markedness are not randomly working in a language, but that their influence on the morphosyntax is regulated through the interaction with other constraints, accounting for systematic differences between languages.

22.4 Generic reference and nominal structure

In this section, we discuss a possible extension of the analysis developed in section 22.2 to plural nominals conveying generic reference. In languages that allow an extensive use of bare nominals, genericity is frequently conveyed by bare nominals (e.g., Mandarin Chinese, Polish, Hindi, etc.). In languages with elaborate plural morphology and a contrast between definites and indefinites, we observe a striking instance of cross-linguistic variation in the expression of plural genericity. In this section, we will be concerned with languages that rank FPL, FDR and FDEF above *FUNCTN in the grammar, and which have strong plural morphology, so that definite plurals contrast with bare plurals.[8] When it comes to the expression of genericity, this class of languages gives rise to two subclasses. One group of languages uses definite generic plurals (Italian, Greek, Hungarian), the other uses bare generic plurals (English, Dutch). Both bare and definite plural generics are found in German, which constitutes an intermediate case.

In line with the literature, in particular Krifka et al. (1995), we distinguish between two types of genericity, namely reference to kinds and generic generalizations. The examples in (16a) and (16b) illustrate that English uses the bare plural in both environments:

(16) a. Dinosaurs are extinct. [reference to kinds]
b. Dogs are dangerous when they are hungry. [generic generalization]

(16a) involves direct reference to kinds, for it is the kind that it extinct. (16b) involves a generic generalization over individual dogs instantiating the kind. Dutch is a language that patterns very similarly to English.

Interestingly, Romance languages such as Spanish, Italian, and Portuguese use definites, rather than bare plurals in both generic environments, as discussed by Longobardi (2001) and illustrated with data from Farkas and de Swart (2007):

(17) a. Dinosaurii au dispărut. [Romanian]
Dinosaur.DEF has disappeared.
'Dinosaurs are extinct.'

b. Cîinii sînt intelegenţi.
Dog.DEF are intelligent.
'Dogs are intelligent.'

This is not a typically Romance pattern, though, for Greek and Hungarian also use definite rather than bare plurals to convey genericity, as shown by Farkas and de Swart (2007):

(18) a. A dinoszauruszok kihaltak. [Hungarian]
the dinosaur.PL die.out.PAST.PL
'Dinosaurs are extinct.'

b. A kutyák veszélyesek mikor éhesek.
the dog.PL dangerous.PL when hungry.PL
'Dogs are dangerous when they are hungry.'

In most languages, there is one designated form for plural generics, either the bare plural, or the definite plural. However, some languages allow both, for instance German:

8 The restriction to languages with strong plural morphology simplifies the discussion because it allows us to concentrate on bare plurals vs. definite plurals. It is not crucial though, for Farkas and de Swart (2007) show that French (which has weak plural morphology) behaves like Italian, Spanish, and Romanian in using definite plurals to convey genericity.

(19) a. Die Dinosaurier sind vor ungefähr 65 Millionen Jahren
the dinosaurs are before about 65 millions years
ausgestorben [German]
died.out
'Dinosaurs died out about 65 million years ago.'

b. Man weiβ nicht genau, warum Dinosaurier ausgestorben sind.
One knows not exactly why dinosaurs died.out are
'We do not quite know why dinosaurs died out.'

The OT typology developed in section 22.2 does not account for the contrast between English and Dutch, on the one hand, Italian, Greek, and Hungarian, on the other hand, nor for the optionality of the definite article in German plural generics, because all these languages adopt the grammar {Fdef, Fdr, Fpl} ≫ *FunctN. In particular, we notice that Fdef is not decisive in generic environments. Farkas and de Swart (2007, 2009) argue that this constraint is independently satisfied by the generic construction in all the languages involved. Accordingly, we need a more fine-grained distinction between the two groups of languages illustrated in (16)–(18). According to Farkas and de Swart (2007), the semantics of genericity involves two ingredients that impose conflicting demands on the nominal expression. On the one hand, the discourse referent involved in generic plurals is not discourse old, because it ranges over individuals across times and possible worlds. On the other hand, the discourse referent involves in generic plurals is always maximal, because it covers the entire kind, or generalizes over all individuals that satisfy a particular description. The observation that the generic plural is not discourse old (represented as a feature [−fam] on the discourse referent) militates against the use of a definite article, because familiarity is one of the ingredients of definiteness, as we saw in section 22.2.3 above. However, the fact that the generic plural involves maximal reference (represented as a feature [+max] on the discourse referent) would favor the use of a definite article, because uniqueness or maximality is another ingredient of definiteness. These conflicting tendencies are captured by the following two constraints:

- *Def/[–fam]: Avoid non-familiar definites
- MaxMax: reflect maximality features of the referent in the nominal structure

"High familiarity" languages like English and Dutch adopt the ranking *Def/[-fam] ≫ MaxMax, and use bare plurals to convey genericity. "High maximality" languages like the Romance languages, Greek, and Hungarian adopt the ranking MaxMax and use definite generic plurals, as Tableaux 22.15 and 22.16 illustrate.

The input meaning is the same in both tableaux. Along the lines of Krifka et al. (1995), it spells out the semantics of a generic generalization in terms of a generic operator ranging over individual dogs instantiating the kind across times and worlds. In addition, the input specifies the values of the discourse referent concerning the features of maximality and familiarity. The tableaux map this input

Tableau 22.15. "High familiarity" languages use bare plurals to convey genericity (e.g., English)

Meaning Gen_x (Dog(x_pl), Intellig(x_pl)) [+Max] [−Fam]	Form	*DEF/[-FAM]	MAXMAX
☞	Dogs are intelligent		*
	The dogs are intelligent	*	

Tableau 22.16. "High maximality" languages use definite generic plurals (e.g., Romanian)

Meaning Gen_x (Dog(x_pl), Intellig(x_pl)) [+Max] [−Fam]	Form	MAXMAX	*DEF/[-FAM]
	Cîini sînt inteligenţi Dog.PL are intelligent	*	
☞	Cîinii sînt inteligenţi Dog.PL.DEF are intelligent		*

onto a number of candidates that crucially vary in article use. The choice between a definite or a bare plural is dictated by the ranking of the two constraints MAXMAX and *DEF/[–FAM] in the two languages. The same meaning can thus be expressed in two ways, and languages differ in the optimal form they associate with the mixed input of familiarity and maximality that is characteristic of plural generics in generic generalizations. Tableaux 22.15 and 22.16 spell out the optimization process for generic generalizations. The case of plural generic subjects of kind level predicates is similar, because they also involve the feature combination [+ max][−fam].

The optional use of definite articles in the German expression of kind reference illustrated in (19) can be approached in different ways. Either the choice for a definite or a bare plural is bound to a particular regional or dialectal version of German, in which case we are dealing with an instance of micro-variation. We can then postulate two varieties of German, call them German_*bare* and German_*def*, which have the rankings in Tableaux 22.15 and 22.16 respectively. Alternatively, the alternation could be due to a change in progress, the two versions reflecting an "older" and a "newer" pattern, where the ranking is dependent on the generation of the speaker, or the register of use. Finally, it is possible that the choice between the two forms is free, which could be reflected in a grammar that ranks the two constraints at the same level, as indicated by the dotted line between the two columns in Tableau 22.17.

The decision between the three options requires a more extensive dataset of German dialects and registers of use than we can investigate in this chapter.

Tableau 22.17. Optional use of a definite article in plural generics (e.g., German)

Meaning Dinosaur(x_{kpl}) & Extinct(x_{kpl}) [+Max] [−Fam]	Form	*DEF/[-FAM]	MAXMAX
☞	Dinosaurier sind ausgestorben Dinosaurs are extinct		*
☞	Die Dinosaurier sind ausgestorben The dinosaurs are extinct	*	

Given that MAXMAX and *DEF/[−FAM] target highly specific ingredients of the semantics of plural definiteness, they immediately become inactive when FDEF and FPL are ranked below *FUNCTN. As a result, such constraints create subclasses of one of the language groups distinguished in Table 22.1 in section 22.2.4, but they do not complicate the entire typology.

What we can learn from the analysis of cross-linguistic variation in the expression of plural genericity is that the course-grained typological classifications defined in section 22.2 can be refined by looking at subclasses of languages. Thus OT tools can be applied in macro-level typological variation, as well as in meso-level variation (within a family of languages), or even micro-variation (across dialects or diachronic stages of a language). Given that typology and language change are mirror images of each other in space and time, we can also use OT to trace diachronic developments in language, as illustrated by Jäger and Rosenbach (2006).

22.5 CONCLUSION

As our point of departure in this chapter, we took the optimization approach developed in Optimality Theory, and applied it to the typology of number morphology and article use. The OT approach provides an alternative to the semantic parameter based theory proposed by Chierchia (1998). Chierchia distinguishes three types of languages on the basis of how common nouns behave as arguments or predicates (proper names are different and will be kept apart here, as in most treatments). In one class of languages (of which Mandarin Chinese is the typical example) every noun is an argument, which means that it does not need article or number marking at all. In another class of languages (of which the Romance languages are the typical examples), every noun is a predicate. This means that a noun can never be used as a subject or object without the support of some sort of marking that turns this noun into an argument. This marking can be an article but it can also be number marking. Then there are languages that have properties of both. They behave like Mandarin Chinese for one class of nouns and like Romance for another class of nouns. English is an example of such a language. Some nouns can occur without any marking (the mass nouns, (20a)), other nouns need marking by an article or plural suffix (in order to be able to occur in argument position, (20b) and (20b′)):

(20) a. I read literature.
b. *I read book.
b′. I read <u>a</u> book/I read <u>the</u> book/I read book<u>s</u>.

In this class of languages the count nouns are predicates (like Romance, in need of marking), the mass nouns are arguments (like Chinese, not in need of marking).

What we see here then is a rudimentary scale ranging from languages that have no grammatical marking on nouns to languages that have obligatory marking on all nouns. In the middle are languages that mark only part of the nouns, see Table 22.2. Languages that are both [−argument] and [−predicate] are ruled out because with such a combination nouns would not be able to occur at all.

There are two important differences between Chierchia's parameter-based approach and the OT analysis advanced here. From the comparison of Tables 22.1 and 22.2 it becomes clear that the OT analysis permits a richer inventory of languages, so it has an advantage over Chierchia's proposal, which has been criticized for lack of empirical coverage (Schmitt and Munn 1999). Within the [+argument, +predicate] languages in the middle of Table 22.2, Chierchia distinguishes between languages that have articles (like the Germanic languages) and languages that don't have articles (like the Slavic languages). However, languages with articles are not all alike. Hebrew, for instance, differs from the Germanic languages in having only a definite article, and no indefinite article. As a result, bare singular nouns do occur in Hebrew (example 4), but not in English (example 5). This requires a finer-grained typology, that goes beyond a simple distinction between languages that do and do not have articles, respectively. In a sense, Hebrew is in between Slavic on the left and Germanic on the right. Chierchia leaves open how this linguistic variety

Table 22.2. Chierchia's (1998) Nominal Mapping Parameter approach

+argument −predicate Chinese	+argument +predicate English	−argument +predicate French

should be characterized, but it is clear that languages are not arbitrary in what articles they possess. When a language has only one type of article, then this article is more likely to be definite than indefinite (Dryer 2007), as we have seen above. Also, within the [+predicate] languages finer distinctions need to made. While French does typically not allow bare nouns at all (in argument position), Italian allows bare plurals and mass nouns in certain argument positions (Longobardi 2001; Chierchia 1998). Both Longobardi and Chierchia explain this difference from a special invisible article in Italian lacking in French. The question is of course why Italian and French would differ in this way and why there is no invisible article for singular count nouns. Although more work needs to be done on the OT typology to provide a more complete description of the rich variation we find in number and article systems across the globe, we think the system proposed in this chapter provides a better starting point for an analysis with a better typological coverage. On the one hand, we can add more constraints to the system developed in section 22.2, as already illustrated in sections 22.3 and 22.4, which allows us to fine-tune the analysis. On the other hand, we observe that constraints interact, whereas the parameters in Table 22.2 are in principle independent of each other. The interaction of constraints is particularly relevant in cases where we observe the "emergence of the unmarked", as pointed out in various places in section 22.2. For example, the account we have developed of obligatory article use with singular (count) nouns in regular argument position (as illustrated in 5a for English, and 6a for St'át'imcets) extends in a natural way to the felicity of bare nominals in non-referential positions (such as 5b for English, and 6b for St'át'imcets) (cf. Tableau 22.6). Such connections are difficult to build into a parameter-based approach.

A second important difference between the approach advanced by Chierchia (1998) and the OT analysis developed in this chapter concerns the locus of cross-linguistic variation. In Chierchia's approach the variation between languages is located in the lexicon. It is a property of nouns in French, for instance, that they are [+predicate], i.e., that they always require marking (in argument positions). Given that the variation is located in the lexicon (it is a property of lexical items), the grammar can be universal. By contrast, OT locates universal grammar in the set of constraints. The constraints discussed in this chapter are claimed to be present in the grammar of all languages. However, whether a constraint is actually operative, i.e., has a visible effect in the language depends on its position in the ranking. Referential faithfulness constraints are thus part of the grammar of languages like Mandarin Chinese and Hindi, but their effect is not visible because they are ranked below the general markedness constraint *FunctN. Given that cross-linguistic variation is accounted for by reranking of constraints, as outlined in section 22.2, OT has means to locate typological variation in the grammar. What we need in linguistics is a system that can deal with sources of cross-linguistic variation at different levels and of different magnitudes that can deal both with the broad typological patterns, but also with the smaller-scale exceptions that we find within languages and within constructions. It should allow us to integrate various factors in one framework so that we can make more precise hypotheses about the way these factors interact. Optimality Theory, as a general framework for formulating hypotheses about constraint satisfaction and interaction, offers new perspectives to achieve this goal.

CHAPTER 23

THE PARALLEL ARCHITECTURE AND ITS PLACE IN COGNITIVE SCIENCE

RAY JACKENDOFF

It has become fashionable recently to speak of linguistic inquiry as *biolinguistics*, an attempt to frame questions of linguistic theory in terms of the place of language in a biological context. The Minimalist Program (Chomsky 1995; 2001*a*) is of course the most prominent stream of research in this paradigm. However, an alternative stream within the paradigm, the Parallel Architecture, has been developing in my own work over the past 30 years; it includes two important subcomponents, Conceptual Structure and Simpler Syntax (Jackendoff 2002; 2007*b*; Culicover and Jackendoff 2005). This chapter will show how the Parallel Architecture is in many ways a more promising realization of biolinguistic goals than the Minimalist Program and that, more than the Minimalist Program, it is conducive to integration with both the rest of linguistic theory and the rest of cognitive science.

23.1 Parallel architectures, broadly conceived

The Parallel Architecture (PA) can be explored at two levels: First, what is a parallel architecture in general? Second, what distinguishes "the" Parallel Architecture from other theories within this genre? In both cases, the basic question is:

(1) What is the best way to allocate the generative capacity of language, so as to account for the observed relations between sound and meaning?

Traditional generative grammar, from *Syntactic Structures* (Chomsky 1957) through the Minimalist Program, has answered:

(2) (Syntactocentric architecture) The recursive rules of the syntactic component provide the generative capacity of language. The relation between sound and meaning arises through mapping syntactic structures into phonetic form (PF) (or the "sensorimotor interface") on one hand and logical form (LF) (or the "conceptual–intentional interface") on the other.

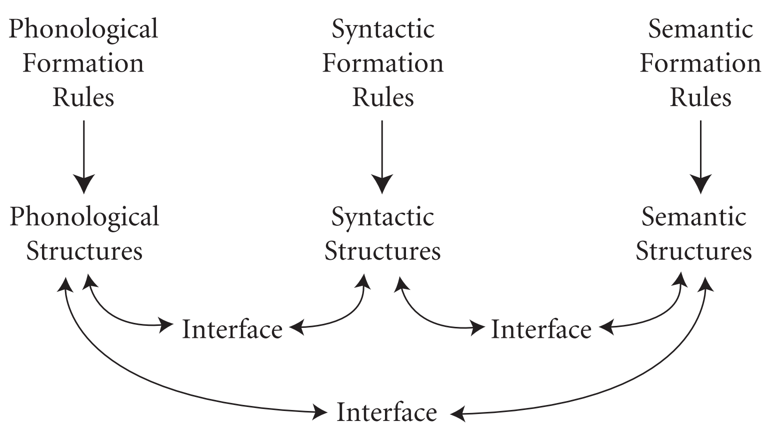
However, theoretical developments as early as the 1970s showed that phonological structures have their own primitives and principles of combination that cannot be reduced to syntactic terms. For instance, rules of syllabification, prosody, and morphophonology are stated in terms of units that are thoroughly comfortable in phonological terms but often quite uncomfortable in syntactic terms. The same is true of meanings: semantic notions like event, manner, quantification, and focus cut across syntactic categories but are characterizable in independent semantic terms. In particular, it has been a staple of linguistic theory and psycholinguistics to distinguish semantic ill-formedness (**Colorless green ideas sleep furiously*) from syntactic ill-formedness (**A good ideas am rare*), which depends on the distinction between semantic and syntactic combinatoriality. (Note: "syntactic" is sometimes used to mean "combinatorial in any sense", including music, phonology, and semantics. I am using the term here in the narrow sense of "combinatoriality whose units are things like Noun and Verb.")

Within the syntactocentric approach, mismatches between phonology and syntax have been either incorporated into syntax (e.g., "Spell-Out") or ignored, at least by syntacticians. More striking has been the constant attempt to build more and more aspects of semantics into syntactic structure—as is required by an architecture in which all combinatoriality is a consequence of syntax. The inevitable result is a syntactic component overstuffed with invisible structure, in which every constituent moves multiple times. Differences of opinion arise as to whether this is a good thing or not.

A parallel architecture answers question (1) like this:

(3) (Parallel architecture)

a. The generative capacity of language is invested in multiple components—at the very least, autonomous generative components for phonological, syntactic, and semantic structure. Each component has its own distinctive primitives and principles of combination, and generates its own structure.
b. The relation between sound and meaning is mediated by a set of interface components, which establish optimal linkings among the various structures and their parts. (Note: in this theory, an interface is not a level of structure but rather a *connection* between *two* levels of structure.)
c. The structure of a sentence is therefore an n-tuple of structures, one for each appropriate component, plus the linkages established among them by the interface components.



A priori, answer (2) seems simpler, since it has has only one "generative engine" and fewer components overall. But, to parallel Chomsky's (1972*b*) rejoinder to Postal's (1972) *The Best Theory*, architectures must be judged not only on their formal elegance but also on their affordance for describing the data of language in full detail (descriptive adequacy), in explaining language variation and the possibility of language acquisition (explanatory adequacy), and in explaining how the system can arise from more general cognitive and biological principles ("beyond explanatory adequacy", to use the term of Chomsky 2001*a*). In particular, formal elegance must not be conflated with biological or cognitive elegance, which might be quite different.

Pursuing the goal of going beyond explanatory adequacy, consider which sort of architecture conforms more closely to what is known about the brain. The visual system is known to contain numerous areas specialized to different aspects of visual perception: detection of motion, detection of color, several independent mechanisms for constructing the perception of depth, possibly face perception, and many

many others. Each of them accounts for a relatively limited aspect of visual understanding; the totality of visual understanding arises from their combined efforts. In order for their efforts to combine, they must communicate, linking their respective representations in an optimal fashion (Koch 2004). At the moment, we don't know a lot about the formal details of representations computed by various visual areas, and there is still much dispute about what brain areas are responsible for different aspects of linguistic understanding. Nevertheless, the overall flavor of the visual system is far more compatible with a parallel architecture, with its multiple independent but linked components, than with a syntactocentric one. No one to my knowledge has ever proposed a visual counterpart of a syntactocentric grammar.

There is one cognitive capacity other than language for which formal details of the representations have been explored in some detail: music. Here it proves impossible to generate musical structures from a single component. Lerdahl and Jackendoff 1983 (see also Jackendoff and Lerdahl 2006) develop a parallel architecture for music containing four components linked by interface rules. One of these structures, grouping, is a general-purpose cognitive capacity that also plays an important role in vision. Another, metrical structure, bears strong similarities to the metrical systems that determine stress and prosody in language. The other two structures are, so far as we know at the moment, particular to music.

One of my original motivations for a parallel architecture in language (Jackendoff 1997, 2002) was the existence of multiple independent tiers in phonology, such as syllabic structure, metrical structure, prosody, and tone, also linked by correspondence or interface rules. Similarly, it is now fairly clear that semantic structure can be dissected into semi-independent structures—at least propositional structure (who did what to whom) and information structure (topic vs. focus vs. common ground). Finally, the relation of language to vision, such that we can talk about what we see, has to be mediated by a set of principles that link linguistic representations of some level to visual representations of some level—it cannot be accounted for through further derivation from syntactic structure (Jackendoff 1987). Thus the internal structure of some components of language, as well as the relation of language to other faculties, is consonant with a parallel architecture for language as a whole.

A parallel architecture for language and other cognitive faculties displays a version of modularity. This is not modularity in Fodor's (1983) sense, which seals off various capacities from each other, but what could be called *representational* or *structure-based modularity*. Each separate form of representation has its own particular autonomous (i.e., domain-specific) structure and its own interfaces to other structures. One form of representation is *relatively* informationally encapsulated from another to the degree that one can influence the other only through a series of interfaces, or through a narrowly specialized interface. For example, phonological structure is relatively encapsulated from visual representations because, in order to speak about what one sees, one has to pass from high-level visual understanding through linguistic semantic structure and syntactic structure in order to influence phonology—i.e., through a series of interfaces. However, there is also a narrowly circumscribed vision-to-phonology interface that subserves reading, and this aspect of vision is rather tightly yoked to phonology. (For more detail, see Jackendoff 1987, chapter 12; 2002, section 7.5.)

In short, the spirit of parallel architectures is in overall accord with what is known about (a) the brain, (b) the structure of other cognitive capacities, (c) the interior structure of linguistic components, and (d) the interaction of language with other cognitive capacities. The syntactocentric architecture, including the Minimalist Program as one realization, is not. (An advocate of Minimalism might respond that this issue is one of performance or implementation, and so this sort of evidence is not pertinent to Minimalist inquiry. I would consider such a response simply a rhetorical avoidance of the evidence.)

Many different theories of grammar employ parallel architectures in this broad sense. As noted above, phonological theory since the mid-1970s has been thoroughly parallel in conception. Among syntactic theories, the most prominent parallel architecture is Lexical-Functional Grammar (Bresnan 2001), where the work of syntax is divided between f-structure, c-structure, and the interface between them. Autolexical Syntax (Sadock 1991) has parallel components for morphosyntactic structure and phrasal syntactic structure, with the possibility of further subdivision. Role and Reference Grammar (Van Valin and LaPolla 1997) subdivides syntax into morphosyntax and phrasal syntax, and semantics into propositional and information structures, with interfaces running in all directions. Construction Grammar (Fillmore 1988, Goldberg 1995) is not formally laid out as a parallel architecture, but it acknowledges the independence of semantics from syntactic form, in that it emphasizes the many–many mapping between syntactic form and meaning, possible only if semantics is autonomous. And the granddaddy of them all is Stratificational Grammar (Lamb 1966), which decomposes the entire grammar into a long sequence of autonomous levels linked by interface components.

Another fundamental question in the architecture of grammar is this:

(4) What formal operations are employed in building linguistic structure?

The mainstream architecture (along with Tree-Adjoining Grammar, Joshi 1987) gives the following answer:

(5) (Derivation-based generation) Syntactic trees are built algorithmically, either from the top down (as in pre-Minimalist theories) or from the bottom up (as in MP and TAG), and they undergo a sequence of distortions (movements and deletions) to derive sound and meaning.

In parallel architectures, the interface relation between different components cannot be a sequenced *derivation*, since structures in different components often stand in a many-to-many relation. Rather, the interface components must be treated as *constraints* (possibly violable), which establish (or license) well-formed links among different kinds of structure. In principle, the rules responsible for each individual component of structure could be algorithmic. But in practice, almost all parallel architectures I have encountered have utilized a constraint-based formalism, in which each independent structure is licensed by simultaneously applied constraints. (An exception is Synchronous TAG, Shieber and Schabes 1991.) To sum up, the answer to question (4) is (6).

(6) (Constraint-based generation) The structures of each component are licensed by simultaneously applied component-internal constraints. The relationships among structures in different components are licensed by interface constraints.

Thus a parallel derivation has no notion of logical sequence, as is essential in a syntactocentric derivation. This has consequences for the relation of linguistic theory to theories of processing, as we will see in the next section.

23.2 The Parallel Architecture: The lexicon

Having settled on a parallel architecture, the more specific question is: What are the autonomous representational formats, and what are the interfaces among them? What I have been calling "the" Parallel Architecture (in capitals, or PA) incorporates specific proposals about semantics, phrasal syntax, and the interface between them, plus less specific proposals about morphology and phonology.

A leading question in the Parallel Architecture is the structure of the lexicon. The question is stated in essentially psycholinguistic terms:

(7) What linguistic material does a speaker have to store in memory—i.e., What is in the lexicon? What structures can be built online in the course of speaking and understanding?

Traditionally, the lexicon is thought of as consisting of words (or morphemes), a distinct component of the language from the rules of grammar. Thinking in terms of question (7) leads to quite a different conception.

A typical word—in any theory—is a triple of phonological, syntactic, and semantic information. In syntactocentric theories, a word is inserted into a syntactic derivation (by lexical insertion or Merge), and it is carried through the derivation to the points where its phonological and semantic properties are "read off". In the Parallel Architecture, the picture is quite different. The structure of a word suits it perfectly to function as a part of the interface components: it establishes a correspondence between small chunks of phonological, syntactic, and semantic structures. (Larger chunks are connected by other interface rules.)

There is no "point in the derivation" where a word is inserted. Rather, one can think of the word being "inserted" into all three structures at the same time, along with the links among them. Or one can think of the word as licensing the connection among preexisting structures. Alternatively, one can think in terms of processing. Given a perceived phonological structure, the word licenses the building of a connection to the corresponding pieces of syntactic and semantic structure; given a piece of meaning to be expressed, the word licenses connecting it to appropriate pieces of syntactic and phonological structures. This last view suits PA to serve directly as a component of a theory of sentence processing (Jackendoff 2002, chapter 7; 2007*a*). PA itself is nondirectional, but its constraints can be implemented in any order suited to particular processing tasks.

Among the information coded in a lexical item is its contextual restrictions. Syntactic contextual restrictions include subcategorization features on syntactic arguments; semantic contextual restrictions include selectional restrictions on semantic arguments. Often these two are partly redundant with each other, but not always (see Jackendoff 2002, section 5.9).

Not every word has to connect all three components. English contains a small collection of "defective" words such as (8a). These have phonology and meaning but no syntactic properties that allow them to combine into larger phrases (aside from within direct quotes, where anything at all is allowed). There are also a few words that have phonological and syntactic properties but no meaning, such as (8b).

(8) a. Phonology and meaning, no syntax
hello, ouch, upsy-daisy, allakazam, wow, shhh, ...

b. Phonology and syntax, no meaning
do (*do*-support), it (pleonastic), of (N *of* NP)

A lexicon conceived in terms of question (7) must contain more than single words. Most obviously, it must contain the thousands of idioms and other fixed expressions in the language such as (9), all of which are units known by native speakers.

(9) a. Idioms
kick the bucket, a breath of fresh air, right on the money, the jig is up, day in day out, clean as a whistle, pie in the sky, ...

b. Fixed expressions (clichés, etc.)
baby-blue eyes, home sweet home, take it from me, weapons of mass destruction, no money down, leave a message at the tone, ...

Including these items in the lexicon (as they must be—where else would they be in the language?) leads to two important conclusions.

First, lexical items cannot be conceived of as syntactic atoms, since many items in (9) have internal syntactic structure. *Kick the bucket* is a transitive VP, *clean as a whistle* is an NP with a comparative complement, *weapons of mass destruction* is a complex NP, and so on. Thus they cannot be inserted by a process like MP's Merge, which builds structure out of syntactic atoms. However, treated as interface constraints, they pose no problem: they simply link a complex syntactic structure with an idiosyncratic meaning. (This approach is shared with HPSG.)

Second, the lexicon cannot be conceived of as a nonredundant list of exceptions, as Chomsky has often asserted (citing Bloomfield). The lexical item *weapons of mass destruction* contains four independently attested words, meaning exactly what they ought to mean. It adds the information that these four form a known unit, and adds some extra meaning or connotation. It is impossible to extract the redundant information, leaving only the extra information, and end up with something that is formally coherent. The conclusion is that the lexicon is full of redundancy. In terms of formal elegance this is less than satisfactory, but it is where the facts urge us. In terms of "brain" elegance, though, it seems entirely in line with the rest of the brain, which seems to favor redundancy where possible, in the interests of more reliable memory and processing.

In addition to items such as (9) that are larger than a word, the PA's lexicon also contains items that are smaller than a word. For example, the regular plural suffix *-z/-s/-əz* in English establishes a correspondence between a piece of phonology, a syntactic feature, and a piece of meaning. Its contextual restrictions state that it is to be affixed to a noun (syntactic context) that is count (semantic context); the conditions for its allomorphy depend on its phonological context. It can be affixed to a noun of any phonological shape, including novel ones (as in the *wugs* test). Thus its manner of combining with its host is formally no different from the way a transitive verb combines with its object, except that it combines below the word level rather than at the phrasal level.

On the other hand, irregular plurals (*oxen*, *women*, *axes*, etc.) have to be learned individually and therefore have to be stored in the lexicon. Formally, they are semantically and syntactically composite, but phonologically unitary. They are therefore parallel in structure to idioms, which are phonologically and syntactically composite but semantically unitary. We can therefore think of these cases as "morphological idioms". (There may of course be subregularities among irregular forms, but we set this aside for purposes of the present chapter; see Jackendoff 2002, sections 6.2–6.4.)

The treatment of regular inflectional morphology as lexical items extends easily to other regular morphological phenomena, including unusual ones. For instance, English expletive infixation (*manu-fuckin-facturer*) is a stored morpheme with a distinct (non-truth-conditional) meaning, and it can be affixed to any syntactic category. Its main contextual restriction is prosodic. Similarly, reduplicative morphemes have meanings and syntactic contextual restrictions just like any other affix, but their phonological shape is listed in the lexicon as a sort of binding: "Copy such-and-such a part of the word I'm attached to" (Ghomeshi et al. 2004).

PA's treatment of regular morphology parts company here with "lexicalist" theories such as LFG and HPSG, which derive morphologically complex words "in the lexicon", "prior to" inserting them into sentences. In PA, both phrasal grammar and morphology contain processes of free combination that can be used online, and both also include lexically listed "prefabs" (idioms and irregular morphological combinations respectively). The difference between phrasal grammar and morphology is only that the units and principles of combination for phrases are in part different from those for words. In this framework, LFG's notion of Lexical Integrity amounts to the claim that the two sets of principles do not interact, except through inflectional morphology.

PA's lexicon also incorporates the insight of Construction Grammar that certain pieces of syntax can carry idiomatic meaning, with or without overt morphemes that mark the constructional meaning. Some of these constructional idioms have ordinary syntax, for instance the VP constructions in (10); others, such as (11), have unusual syntax ("syntactic nuts" in the sense of Culicover 1999).

(10) a. joke your way into the meeting (V Pro's way PP = 'go PP while/by V-ing')
b. rumble around the corner (V PP = 'go PP in such a way to make a V-ing sound')
c. knit the afternoon away (V NP away = 'spend NP[time] V-ing')
d. paint me a picture (V NP_1 NP_2 = 'V NP_2 for the benefit of NP_1')

(11) a. The more you eat, the fatter you get (the more S, the more S)
b. One more beer and I'm leaving (one more X and S)
c. student after student (N P N)
d. How about some lunch? (How about XP?)

Each of these constructions is listed in the lexicon as a linking between a syntactic complex and a meaning; some parts of the syntactic complex may be linked also to phonology (e.g., *way*). The syntactic variables in these constructions correspond to semantic variables in the usual way, and the constructions can therefore be combined with other items to form a sentence in exactly the same way as words and other idioms are. (However, notice that the verbs in (10a, b, c), though they are syntactic heads, serve semantically as manner or means modifiers.)

Since the lexicon contains linked phonological, syntactic, and semantic complexes, nothing in principle prevents it from also containing phonological and syntactic complexes that are not inherently linked to anything. For example, a "generative" phrase structure rule such as (12a)—which, as part of one's knowledge of English, must be stored in memory somehow—can also be stated as a "treelet" (12b), a syntactic complex that constrains possible syntactic structures. PA treats it as a stored piece of structure; it can therefore be localized in the lexicon alongside semantically and phonologically linked VPs such as *kick the bucket*.

(12) a. VP → V – NP
b. $[_{\text{VP}}$ V NP]

Thus, to the extent that there are autonomous principles of syntax such as fixed head position, the availability of ditransitive constructions, the means for forming relative clauses, and so on, these are stated in precisely the same format as constructional idioms, and they therefore belong in the lexicon as well. In phonology, one can view syllable structure rules as lexical entries that specify pieces of autonomous phonology.

The upshot is that there is no principled distinction between words and rules of grammar. Both are stored pieces of structure, lying at opposite ends of a multidimensional continuum of idiosyncrasy and regularity. This conclusion has been arrived at within HPSG (Pollard and Sag 1994), Cognitive Grammar (Langacker 1987*b*), and Construction Grammar as well as PA, in each case through attention to a multitude of intermediate cases such as idioms and constructions. Mainstream generative grammar, partly because of its algorithmic formulation, has followed traditional grammar in making a strong lexicon/grammar distinction. This has made it difficult to assimilate idioms and constructions into the theory, resulting in loss of descriptive adequacy.

In pursuit of explanatory adequacy, the MP has arrived at the conjecture that there is actually only one rule of grammar, Merge, and that all differences among languages are localized in the lexicon (Chomsky 2001*a*); this conjecture has not proven as simple in execution as in principle (particularly since MP has *no* theory of the organization of the lexicon!). Within PA, HPSG, and Construction Grammar, the counterpart of this conjecture is quite straightforward. All words and all rules of grammar are pieces of structure stored in the lexicon. The only "procedural" part of language is the fundamental operation of Unification (Shieber 1986), which assembles pieces of structure. Merge proves to be a special case of Unification: it combines two given elements with a piece of tree structure.

Unification can be generalized to combinatorial cognitive capacities other than language, thus better satisfying the goal of "beyond explanatory adequacy". For example, in vision it can be used to integrate evidence for depth perception from disparate sources. It can also be used to weld lyrics to music in building songs. Merge cannot perform either of these functions. If Unification is a general brain mechanism for achieving combinatoriality, it should be no surprise that language uses it too. (See Jackendoff 2008 for discussion of Merge vs. Unification.)

23.3 Conceptual Semantics

To work out any version of a parallel architecture, it is necessary to have theories of the individual components and the interfaces among them. Unlike other parallel architectures in the literature, and unlike mainstream linguistic theory, PA is grounded in a highly articulated theory of semantics, Conceptual Semantics, that answers to the concerns of the biolinguistic perspective and that also offers considerable (and continually increasing) empirical coverage. There is space here only to list some of the highlights of the theory.

First, Conceptual Semantics (like Cognitive Grammar) is thoroughly mentalistic: it is a theory of the information in a language user's mind/brain that is involved in understanding utterances, connecting them to perceptual evidence, and making inferences. It recasts the traditional philosophical concerns with reference and truth in mentalistic terms:

(13) a. Traditional formulation:
  i. A phrase P refers to an entity E in the world (or in a possible world).
  ii. A sentence S is true if it meets conditions $C_1, \ldots, C_n$ in the world.

 b. Mentalistic formulation:
  i. A language user LU understands a phrase P to refer to an entity E in the world as LU conceptualizes it.
  ii. LU judges a sentence S true if S meets conditions $C_1, \ldots, C_n$ in the world as LU conceptualizes it.

The seeming objectivity of language, stressed by traditional philosophy of language, is a consequence of language users sharing a common (or near-common) conceptualization of the world, so that agreement can largely be taken for granted (Jackendoff 1983; 2002, chapters 9 and 10).

Second, Conceptual Semantics recognizes that many aspects of one's conceptualization of the world are independent of language. For instance, one can understand much of the behavior of physical objects ("naive physics") without any language at all. Decades of research on child development, linguistic and nonlinguistic, have shown that prelinguistic children bring a rich toolkit to the task of understanding the physical world, and that this understanding serves as a foundation for learning word meanings (e.g., solving Quine's *gavagai* problem). Thus the view of meaning espoused by Conceptual Semantics offers the potential of explanatory adequacy, i.e., helping to explain the innate basis from which children acquire lexicons (now including rules of grammar).

It also appears that other primates—especially apes—negotiate the physical world in much the same way we do; humans differ only in being able to talk about it. This provides an evolutionary underpinning for the semantic system of language: our ancestors had thoughts—as it were, things to talk about—before they could talk. This view of meaning, then, helps satisfy the goal of "beyond explanatory" adequacy: it helps explain why (some part of) the semantic system of language is the way it is, because it is built upon pre-existing primate cognition.

Within the MP, by contrast, the combinatorial properties of the "conceptual–intentional interface" arise through derivation from the syntactic component. On the face of it, this amounts to the claim that babies and apes cannot think combinatorially. It is possible to read certain passages of Chomsky as endorsing such a claim, but to my knowledge it has not been defended against the copious literature on primate intelligence. In a recent passage, Chomsky (2006) says "unbounded Merge provides only a language of thought, and the basis for ancillary processes of externalization". In a way this acknowledges the combinatorial character of thought, but it still does so in syntactocentric terms: the basic units of his "language of thought" are NPs and VPs; and Merge, the capacity for combinatoriality, is said to have arisen in the course of human evolutionary divergence from other primates.

In PA, by contrast, the "language of thought" is the combinatorial system in terms of which one understands the world. Its units are entities such as objects, events, properties, and trajectories. NPs and VPs are part of the combinatorial system of (narrow) syntax, which plays a role in the mediation between thought and sound, that is, as part of what Chomsky calls "processes of externalization". PA takes the combinatorial system of meanings to be universal (though use of the system can be biased by the means of expression, if "Whorfian" effects prove to be genuine). It is just that meanings are not made of syntactic units. This approach is possible precisely because of the fundamental assumption of PA that language—and the mind in general—utilizes multiple sources of combinatoriality.

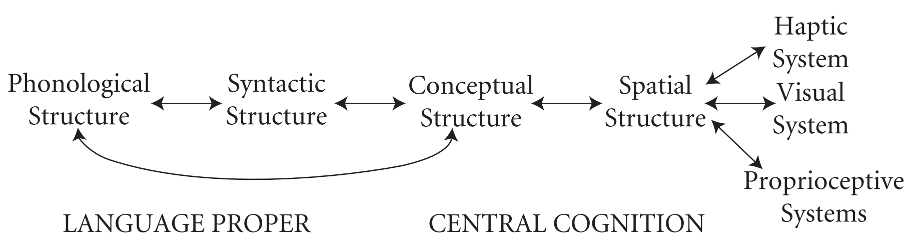
A third important aspect of Conceptual Semantics, again drawing on the Parallel Architecture, is that the system of meaning or "language of thought" is itself bifurcated into two linked combinatorial systems (at least). One of these, Spatial Structure, is quasi-geometric or topological in character. For a first approximation, it might be thought of as the highest level of the visual system. At this level, objects can be represented in terms of their detailed shape. However, shapes are encoded in a perspective-independent fashion, so that they can be recognized from any angle. Objects can also be represented schematically, so that, say, the action of sitting can be represented in terms of a generic or schematic human figure rather than a specific person.

In fact, though, Spatial Structure is not exclusively visual: it can also code shape and configuration that has been derived haptically (sense of touch) and proprioceptively (body sense), and both of these can be compared and integrated with visual input. Thus Spatial Structure is more abstract and general than a visual image—it is conceived of as a central level of cognition that codes the physical world in a relatively modality-independent fashion.

The second major division of meaning is Conceptual Structure, an algebraic structure built up in terms of discrete features and functions. It encodes distinctions that cannot be represented in the geometric/topological format of Spatial Structure, such as those in (14).

(14) a. the type-token distinction, distinguishing categories from individuals
b. taxonomic relations: 'X is an instance/subtype of Y'
c. temporal relations: 'X is past/future'
d. causal relations: 'X causes Y', 'X enables Y', 'X impedes Y', ...
e. modal notions: 'X is hypothetical/nonspecific/potential/fictional ...'
f. social notions: 'X is the name of Y', 'X is dominant to Y', 'X is kin to/friend of Y', 'X is member of group Z', 'X owns Y', 'X is obligated to perform act Y', 'action Y is of normative value to X', ...
g. theory of mind notions: 'X believes Y', 'X imagines Y', 'X intends Y', 'X is committed to norm Y', ...

The overall architecture looks like this:



Conceptual Semantics takes it that word meanings must be composite in order to encode relations among word meanings and in order to state properly general rules of inference. On the other hand, it differs from classical views of word meaning in admitting conditions other than necessary and sufficient. For instance, the conditions for color words must be encoded in terms of relative distance from central instances. In judging a hue between focal red and focal orange, two such conditions come into competition, and the judgment is therefore variable and to some degree context-dependent.

In addition, many word meanings contain multiple conditions interacting in "preference rule" fashion. For instance, sterotypical *climbing* involves moving (i) upward, (ii) in a clambering fashion. But one can *climb down a tree* (clambering but not moving upward), and an airplane can *climb into the clouds* (moving upward but not clambering). On the other hand, an airplane cannot *climb down out of the clouds*, because such motion is neither upward nor clambering. In other words, neither condition is necessary, either may be sufficient, and stereotypical cases satisfy both. This type of rule interaction produces so-called "cluster concepts", of which Wittgenstein's (1953) example of *game* is the most famous.

These characteristics of word meanings, even if strange according to standard philosophical preconceptions, are totally normal within the context of brain computation. As has been observed since the gestalt psychologists of the 1920s (Wertheimer 1923), conditions based on central instances and rule interactions with the characteristics of preference rules are standard in vision. They also appear in phonetic perception and in musical cognition, and essentially anyplace that multiple factors can either combine or conflict in producing a judgment.

Conceptual Semantics differs from most theories of semantics (but again, not from Cognitive Grammar) in that it denies a sharp division between linguistic meaning and encyclopedic meaning (or "knowledge of the world"). Every division that has been proposed turns out to eviscerate linguistic meaning to the point where it cannot serve as a basis for inference (see Jackendoff 2002, sections 9.6–9.7, as well as Bolinger 1965, Langacker 1987*b*, and Levinson 2000).

A related point is that "semantics" and "pragmatics" do not involve distinct representations. Rather, there is a pair of mental representations, Conceptual Structure and Spatial Structure, that are the locus of sentence understanding. Some parts of these representations may come from the words in the sentence and their grammatical configuration; we may call these parts "semantic". Other parts come from nonlinguistic sources such as perception, inference, and "world knowledge"; we may call these parts "pragmatic." But these parts are often intricately interwoven in the representation in such a way that one cannot do the "semantics" first and paste in "pragmatics" afterward.

In Conceptual Semantics, the taxonomy of concepts ("a poodle is a kind of dog", "a dog is a kind of animal", etc.) grounds out in a fundamental ontology of concepts—the basic types of things that humans can conceptualize in the world. Traditional philosophy of language and formal semantics attempt to make do with an absolutely minimal ontology such as individuals and truth-values. Perhaps this makes sense if one thinks semantics is about the nature of reality and should ground out elegantly in fundamental physics. But if semantics is about the human conceptualization of the world, its fundamental units are the product of evolution building a brain equipped to guide an organism successfully through its life. Again "brain elegance" takes precedence over formal elegance.

One piece of evidence for the basic ontology comes from deictic expressions that pick out units in the visual field. Just as it is possible to point out objects for the hearer to identify, as in (15a), it is possible to pick out a wide range of other entities.

(15)	a. Please pick that [pointing] up.	[object]
	b. Please put your hat here [pointing].	[location]
	c. He went thataway [pointing].	[path/trajectory]
	d. Please don't do that [pointing] around here any more.	[action]

e.	Did you hear that?	[sound]
f.	I hope that [pointing] doesn't happen again.	[event]
g.	The fish I caught was this long [demonstrating].	[distance]
h.	There were about this many [gesturing] people here last night.	[amount/number]
i.	Can you walk like this [demonstrating]?	[manner]

Each of these ontological categories has its own conditions of individuation; many of them (but not all) allow a type-token distinction; many permit quantification. Adopting this relatively rich system from the start affords Conceptual Semantics a broad descriptive capacity and, to some extent, a better constrained relation between semantic and syntactic categories. Note also that (15) lists only ontological categories observable in the physical world; there are clearly others, such as information and value.

Once the ontological system is laid out, it becomes possible to recognize entities that subsist simultaneously in more than one ontological domain (the "dot-objects" of Pustejovsky 1995). For instance, a *book* is simultaneously a physical object and a body of information. These two characterizations, moreover, are in a preference rule relation, since there are blank (i.e., informationless) books and books stored on a computer (i.e., not laid out on paper pages). *Reading* is a "dot-action", in that it involves both the physical act of scanning the page with one's eyes and the informational act of receiving information off the page. Dot-objects are therefore multidimensional entities within Conceptual Structure.

Perhaps the most important case of a dot-object is a human being, who is conceptualized simultaneously as an animate physical object and as a person—an entity in the social domain. The two domains correspond to the (nearly universal) cultural conceptualization of people as composed of body and mind (or soul or spirit). The fact that people have faces and hands and livers falls into the physical domain; the social notions and theory-of-mind notions in (14f,g) above are predicated in the social domain. Again, in traditional beliefs at least, these two characterizations stand in a preference rule relation. For instance, a *zombie* is an animate physical object lacking conscious personhood; a *ghost* is a mind (or soul) lacking a physical body. Reincarnation and body-switching (both amply attested in human narratives) are one mind inhabiting different bodies in succession; multiple personality disorder is experienced as different personalities (i.e., different individuals) inhabiting the same body in succession (Jackendoff 2007*b*, chapter 5).

The combinatorial possibilities of Conceptual Structure arise from (at least) three principles of combination: argument satisfaction, modification, and binding. In the default case, argument satisfaction is expressed by syntactic complementation, and modification by syntactic adjuncts. For instance, in *John slept along the river*, *John* expresses an argument of *sleep*, and *along the river* expresses a place modifier. However, there are exceptions to this typical configuration. For instance, in the sound+motion construction illustrated in (10b) above (e.g., *The trolley rumbled along the river*), the subject is a semantic argument not only of the verb but also of an unexpressed predicate of motion. The PP is also an argument of the predicate of motion, and the verb expresses a modifier of this predicate, i.e., "move *while rumbling*". A mismatch in the opposite direction is illustrated by *Bill buttered the bread with cheap margarine*. Here *cheap margarine* is syntactically an adjunct, but semantically it is an argument: it is what Bill put on the bread. Such mismatches are common.

Binding, a direct connection between one conceptual constituent and another, comes in two varieties: identity of reference and identity of sense. This is reflected in two kinds of anaphoric elements in language. Identity of reference binding is expressed by definite pronouns and also by anaphoric epithets, such as in *John wants to win, but <u>the poor guy</u> never will* (which does not display identity of sense). Identity of sense binding is expressed by *one*-anaphora and also by VP anaphora with expressions like *do so*. These two types of binding must be distinguished in Conceptual Structure since they give rise to different inferences.

Using argument satisfaction to create semantic combinations requires functions whose arguments are to be satisfied. A number of broad families of functions have been investigated within Conceptual Semantics:

- Functions that encode spatial location, motion, and orientation. They all take two arguments: a Theme (the object being located or in motion) and a Location or Path: BE(Theme, Loc), GO(Theme, Path), STAY(Theme, Loc), ORIENT(Theme, Path), EXTEND(Theme, Path).
- Functions that encode Locations and Paths relative to a reference object: IN(X), ON(X), TO(X), FROM(X), TOWARD(X), NEAR(X), etc. Some of these involve imposing a reference frame on the reference object; e.g., BEHIND(X) must be specified as to whether one is speaking of the intrinsic back of X or its other side relative to the speaker. (This family has been heavily investigated within Cognitive Grammar as well.)
- Causative functions that encode a Causer (an Agent or Event) being causally connected to an Effect (another Event): CAUSE(Causer, Effect), LET(Causer, Effect), HELP(Causer, Effect), ENABLE (Causer, Effect), and others.
- Mereological functions that encode part–whole relations: PART-OF (legs, handles, noses), BOUNDARY-OF (edges, surfaces, ends, etc.), MEMBER-OF (members of aggregations), COMPOSED-OF (ingredients of mixtures).

A founding insight of Conceptual Semantics (due to Gruber 1965) is that all of these functions can be applied to semantic fields other than physical space. For instance, an object being owned by someone (a social relation) is often expressed crosslinguistically as the object "being at" the owner, and changes of possession are often expressed as the object "going" "from" the previous owner "to" the new owner. Similarly, just as we talk about the end of a rope, we can talk about the end of a speech, a relationship, or a genealogical line. This suggests that these Conceptual functions can be decoupled from their physical context (where they connect with Spatial Structure) so as to apply to more abstract domains as well. In addition to possession, they also extend to such fields as time, event structure (such as aspectuality and telicity), ascription of properties, and (in the case of causation) social coercion and logical entailment. (This insight is treated somewhat differently in Cognitive Grammar (Lakoff 1987), where it is taken to show that underlying linguistic expression is an extensive and powerful system of conceptual metaphor.)

Further functions that have been investigated (Jackendoff 2007*b*) involve the personal domain. They include:

- Theory-of-mind predicates, e.g., "X perceives Y (in various modalities)", "X is committed to proposition P" (belief), "X is committed to action A" (intention), "X is committed to norm N" (adherence to norms).
- Value predicates in various domains (affective, normative, quality, etc.): "X is of value V", "X is of value V to person Y".
- Predicates of exchange: "X does action A in exchange/return/retaliation for Y doing action B".
- Obligations, rights, and authority: "X is obligated to Y to perform action A", "X has a right to perform action A", "X has authority over Y's performing action A".

All of these functions are involved in constructing the *propositional tier* of Conceptual Structure. In addition, sentence meaning involves an *information structure tier*, which designates certain semantic constituents as topic, certain as focus, and the rest as common ground. Further differentiation of the propositional tier has also been proposed, for which there is no space here: a *referential tier* in Jackendoff 2002 (involved for instance in identity-of-reference anaphora, specificity, referential opacity, and quantification) and an *action tier* or *macrorole tier* in Jackendoff 1990*a*, 2007*b*.

In short, Conceptual Semantics aspires to the formal richness necessary to encode the character of human concepts and their inferential affordances. It integrates comfortably with the Parallel Architecture, in that, although it is a combinatorial system, its units and principles of combination—as well as the resulting structures—are quite different from those of syntax. In particular, it is a multidimensional formal system, in that it involves both Spatial Structure and Conceptual Structure, the latter itself split into multiple tiers connected by interface components. Only through looking at semantics on its own terms, grounded in the character of nonlinguistic cognition, can the independence of these structures from language—and their psychological and biological grounding—be revealed. If meanings have this sort of structure, they certainly cannot be derived from the syntax of NPs and VPs.

23.4 Simpler Syntax and the syntax–semantics interface

An advantage of a parallel architecture over a "single-engine" architecture is that no single level of structure has to carry the entire informational load. In a syntactocentric architecture, all semantic combinatoriality has to be derived from syntactic combinatoriality. Thus syntax is forced to be combinatorially at least as complex as semantics—if not more so, since it also has to answer to its own internal imperatives such as word order and agreement. And indeed this outcome has been achieved twice in the history of generative grammar: the first time, in the Generative Semantics movement of the late 1960s and early 1970s (Lakoff 1971), and the second time, in Government-Binding Theory of the 1990s and the Minimalist Program. In MP, the rules of grammar and the contents of UG have been reduced to a minimum (allegedly—though only through drastic cuts in empirical coverage), but the structures and derivations have increased steadily in size (see Culicover and Jackendoff 2005, chapters 2 and 3).

In PA, the combinatorial properties of meaning are a property of autonomous conceptual combinatoriality. From this perspective, syntax functions in the grammar not as the fundamental generative mechanism but rather as an intermediate stage in the mapping between meaning and sound (in either direction). Words are interface rules that provide small-scale mappings between meaning and sound. What remains to complete the mapping is the relationships among the words: the function–argument and function–modifier relations, as well as binding relations. Syntax can be thought of as a way of recoding the semantic relationships among the words in a phrase or sentence in terms that are visible to phonology, such as linear order, inflectional morphology, and anaphoric elements—as well as coding the overall semantic force of a clause, such as declarative vs. interrogative. However, there is no need for syntax to encode any more of semantic structure than is necessary in order to mediate the mapping between phonology and meaning.

In fact, many aspects of meaning are not supported by syntactic or lexical expression, for instance:

(16) a. Implicature:
Are you going to be going near a mailbox? (= "Will you mail some letters for me?")

b. Ellipsis:
It seems we stood and talked like this before. We looked at each other in the same way then. But I can't remember <u>where or when</u>.
[Spoken to someone about to jump off a building] <u>Don't!!!</u>

c. Constructional meaning:
The trolley rumbled around the corner. (="The trolley went around the corner rumbling") (cf. (10b))

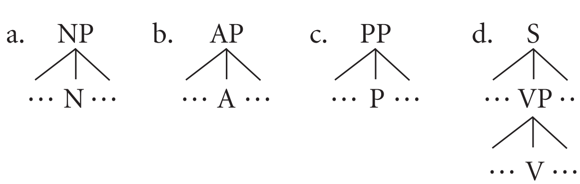
d. Coercion:
The <u>ham sandwich</u> over in the corner wants more coffee. (="guy with ham sandwich")
<u>Plato</u> is on the top shelf. (="book by/bust of Plato")
Joe <u>jumped</u> until the bell rang. (="jumped repeatedly")

Some of these are treated in mainstream theory in terms of syntactic (or PF) deletion of unexpressed elements; others are not treated in mainstream theory at all. Culicover and Jackendoff 2005 show that they are all best treated in terms of elements of semantics that have no syntactic realization.

Carrying this outlook consistently through the syntactic component leads to the approach of Simpler Syntax (Culicover and Jackendoff 2005): an attempt to cut syntactic structure down to the bare minimum necessary to accomplish the sound–meaning mapping. This is a "minimalist" approach to language, but with different premises about what is to be minimized than the Minimalist Program.

The basic stance of Simpler Syntax is that the complexity of semantics is independently necessary in order to explain inference and the relation to perception. Therefore semantics should play as large a role as much as possible in constraining grammaticality, and syntax as little as possible. On the other hand, the "generative engines" of syntax and morphosyntax are still necessary to account for differences among languages in word order, case marking, agreement, handling of long-distance dependencies, and the existence of special constructions. The resulting syntactic theory is by no means simple, but it is far simpl*er* than mainstream models.

The Simpler Syntax lexicon is as described in section 23.2: it contains words, regular affixes, idioms, constructions, and independent principles of phrase structure. Syntactic structures are as flat (i.e., as undifferentiated) as possible. Aside from linear order, there is no *syntactic* distinction between specifiers, arguments, and adjuncts, as this is already provided for in the semantics. The result is predominantly two-layer X-bar skeleta, as in (17a–c). The exception is S, which is a three-layer projection of V, as in (17d).

(17)


One price of this structural simplification is the need for trees with multiple branching nodes rather than strictly binary branching as in MP. Culicover and

Jackendoff 2005 give arguments why strictly binary branching is not an advantage, and in fact is often a disadvantage. Another price of this simplification is that some rules of grammar have to be sensitive to linear order as well as dominance. This is too often taken to be a disadvantage. But from a larger perspective it is actually an advantage. Linear order is given for free in the signal and hierarchical structure is not. So rules that depend in part on linear order ought actually to be easier for the child to learn.

Simpler Syntax makes use of almost no empty nodes in syntactic structure. This is desirable in principle, because empty nodes make heavier demands both on the learner and on processing. Most empty nodes in the classical theory are posited either for semantic reasons or to promote syntactic uniformity. For instance, the phonologically empty element PRO is posited to fill in a semantic subject of an infinitival VP where there is none at the surface, thereby giving all verbs a syntactic subject. Simpler Syntax instead allows infinitival VPs without syntactic subjects, and it uses the interface to identify their "understood" subjects in Conceptual Structure.

Similarly, ellipsis is not derived through empty nodes or deletion. Rather, elliptical configurations, especially when they are syntactically unusual (as in Gapping), are treated as meaningful constructions listed in the lexicon. The interpretation of an elliptical construction is derived from the Conceptual Structure of its antecedent—or from the Conceptual Structure of the context—not from a deleted syntactic structure. Culicover and Jackendoff show many cases of ellipsis for which there is no plausible syntactic antecedent, such as those in (16b).

A standard argument for syntactically derived ellipsis is that elliptical constructions often display syntactic properties that normally can arise only through syntactic licensing (so-called *connectivity*). For instance, in the dialogues in (18), the difference in the prepositions in the replies can be traced directly to the difference between the syntactic licensing of *proud* vs. *pride*.

(18) a. A: Bill is very proud.
B: Yes, especially <u>of</u> his stamp collection. [cf. *proud of/*in*]

b. A: Bill has a lot of pride.
B: Yes, especially <u>in</u> his stamp collection. [cf. *pride in/*of*]

However, similarly licensed syntactic properties appear even in sentences where there is no relevant linguistic context, such as *Do you like these?* [pointing at a pair of pants]. Simpler Syntax proposes a relation of *indirect licensing* that accounts for these effects.

Like other constraint-based theories, Simpler Syntax has no movement and no covert level of syntactic structure such as Logical Form. The effects ascribed to movement in mainstream theory are accounted for with a variety of mechanisms, most of them shared with other constraint-based theories, especially HPSG. These mechanisms include:

- Free phrase order (e.g., among adjuncts in VP, where the order is constrained only by prosody and focus).
- Alternative argument realizations (e.g., dative alternation).
- For long-distance dependencies, operator–trace relations along the lines of HPSG (trace is the only kind of empty node in Simpler Syntax). The constraints on long-distance dependencies arise from multiple sources, only some of which are syntactic. Others arise from processing complexity and from semantics, especially information structure and referential structure.
- Binding and control are relations over Conceptual Structure, not over syntactic structure, though they may involve syntactic conditions on the relation between anaphoric elements and antecedents.

In order to account for so-called A-movements, in particular passive and raising, it is unfortunately necessary to introduce extra machinery. Simpler Syntax proposes a *grammatical function tier* (*GF-tier*) that modulates the syntactic realization of semantic arguments expressed as NPs, that is, subjects, objects, and indirect objects. We are not too dismayed by this extra mechanism, as the principles behind it appear in every substantive syntactic theory: as f-structure in LFG, as essentially *all* of Relational Grammar, as the complement hierarchy in HPSG, and as abstract case in GB/MP.

The analysis is closest to that in LFG and HPSG. However, in these two theories, passive is a rule that converts active verbs into passive verbs in the lexicon, altering their argument structure. As mentioned earlier, this is not an option in PA, where the lexicon is where items are stored, and working memory is where structures are built online. Hence, in Simpler Syntax, passive is treated as a construction that alters argument realization online without altering the verb itself. The GF-tier is of course another piece of parallel architecture, this time a partial mediator of the syntax–semantics interface.

(19) illustrates the linking between the various structures in an example involving raising. The linking relations are notated as subscripts; for visual clarity, some of them are also notated redundantly by vertical association lines.

(19) *John seems to like scotch*:

[SEEM ([LIKE (JOHN$_3$, SCOTCH$_4$)]$_2$)]$_1$	Conceptual Structure
[GF$_3$]$_1$ [GF$_3$ > GF$_4$]$_2$	Grammatical Function Tier
[$_S$ NP$_3$ [$_{VP}$V$_1$ [$_{VP}$ to$_5$ V$_2$ NP$_4$]$_2$]]$_1$	Syntactic Structure
John$_3$ seems$_1$ to$_5$ like$_2$ scotch$_4$	Phonological Structure

In Conceptual Structure, *JOHN* is an argument of *LIKE*. It links to the GF array associated with the subordinate clause (bracketed expression subscripted *2*). In turn, this GF is linked to a GF in the main clause array (subscript *1*), which is then linked to the subject of the main clause and its phonology. The linking through the GF-tier is the Simpler Syntax counterpart of an A-chain in classical syntax. But it is not in syntax proper, as there is no *syntactic* subject at all in the subordinate clause, only a GF-subject. (See Culicover and Jackendoff 2005 for more motivation and detail.)

23.5 Concluding remarks

An abiding issue between linguists and psycholinguists has been the competence–performance distinction. Mainstream linguistics tends to say that the grammar written by linguists is a description of competence, but it is somewhat obscure how it is utilized in performance. This has the effect of insulating linguistic theory from results in psycholinguistics. By contrast, in the Parallel Architecture, language processing consists of assembling pieces of structure stored in the lexicon to form a triple of phonological, syntactic, and semantic structures in working memory. As a result, there is no mystery to the competence–performance distinction. Competence theory describes the pieces of structure and their affordances for assembly, while performance theory describes how these very pieces are assembled in real time, starting from either phonetic input (perception) or conceptual input (production). Details of a performance model in such a vein appear in Jackendoff 2002, chapter 7 and Jackendoff 2007*a*.

The Parallel Architecture also offers an attractive vehicle for discussion of the evolution of the language capacity. It begins with the premise that some version of Conceptual Structure is present in apes, and therefore in our hominid ancestors. Bickerton 1990 and Givón 1979 have proposed that, prior to the development of modern language, there was a stage of "protolanguage", which persists in the human language capacity and emerges in situations such as pidgins and agrammatic aphasia. The defining characteristics of protolanguage are words concatenated into utterances, but lacking any syntactic organization beyond that afforded by linear order. A great deal of the informational load in such an utterance is carried by pragmatics. Within the Parallel Architecture, this form of language can be characterized in terms of a level of phonology linked to Conceptual Structure without the intervention of syntactic structure (Jackendoff 2002, chapter 8).

From this stage, the evolution of a syntactic capacity can be seen as adaptive: it is a canonical coding of semantic relationships among words for greater accuracy and efficiency. In any architecture, phonological and semantic structures have to be relatively rich, as they code the thousands of distinctions among words. In Simpler Syntax, syntactic structure is relatively lean: its elements comprise only a few parts of speech and phrasal categories, as might be expected of a relatively late evolutionary add-on. By contrast, in the mainstream architecture, an elaborate syntax would have had to evolve first before combinatorial phonology and semantics could be possible, a rather less enticing scenario.

To sum up, this chapter has shown many ways in which the Parallel Architecture, with its components Conceptual Semantics and Simpler Syntax, instantiates the biolinguistic outlook better than does the Minimalist Program. In particular, it offers the prospect of integrating linguistics fully with cognitive science. There still remain, of course, many challenges to the approach, among which perhaps the most important are integrating phonology, morphology, language variation, and language change into the model, so that it covers a broader range of linguistic phenomena. In addition, a theory of language acquisition has been sketched (Jackendoff 2002, chapter 6), but it remains a promissory note. It is dearly to be hoped that some of these challenges can be undertaken by practitioners of the relevant subdisciplines.

CHAPTER 24

NEO-GRICEAN PRAGMATIC THEORY OF CONVERSATIONAL IMPLICATURE

YAN HUANG

This chapter is dedicated to Professor P. H. Mathews, my mentor at Cambridge, on the occasion of his seventy-fifth birthday.

In the William James Lectures delivered at Harvard in 1967, H. P. Grice presented a panorama of his thinking on meaning and communication—what he called his "tottering steps" (Grice 1989: 4) toward a systematic, philosophically inspired pragmatic theory of language use, which has since come to be known as the Gricean pragmatic theory (see Chapman 2005 on the life and work of Grice). Since its inception, the classical Gricean paradigm has encouraged numerous refinements, reinterpretations, and reconstructions, giving rise to various neo-Gricean enterprises. Consequently, the classical and neo-Gricean theory has revolutionized pragmatic theorizing and has to date remained one of the two foundation stones of contemporary thinking in linguistic pragmatics and the philosophy of language.

This chapter undertakes to present and assess the neo-Gricean pragmatic theory of conversational implicature. The organization of the essay is as follows. Section 24.1 discusses the classical Gricean theory. Next in section 24.2, I present the neo-Gricean pragmatic theory, focusing on the dualistic model put forward by Horn and the trinitarian model posited by Levinson. Finally, sections 24.3, 24.4, and 24.5 examine the role played by the neo-Gricean pragmatic theory in effecting a radical simplification of lexicon, semantics, and syntax, respectively.

24.1 Classical Gricean theory of conversational implicature

On a general Gricean account of meaning and communication (e.g., Grice 1989), there are two theories: a theory of meaning$_{\text{n[on]n[atural]}}$ and a theory of conversational implicature. In his theory of meaning$_{\text{nn}}$, Grice emphasized the conceptual relation between natural meaning in the external world and non-natural, linguistic meaning of utterances. He developed a reductive analysis of meaning$_{\text{nn}}$ in terms of the speaker's intention, the essence of which is that meaning$_{\text{nn}}$ or speaker-meaning is a matter of expressing and recognizing intention.

In his theory of conversational implicature, Grice suggested that there is an underlying principle that determines the way in which language is used maximally efficiently and effectively to achieve rational interaction in communication. He called this overarching dictum the co-operative principle and subdivided it into nine maxims of conversation classified into four categories: Quality, Quantity, Relation, and Manner. The co-operative principle and its component maxims ensure that in an exchange of conversation, truthfulness, informativeness, relevance, and clarity are aimed at.

(1) Grice's theory of conversational implicature (simplified) (Huang 2000*a*: 206; 2007: 26)

a. The co-operative principle
Be co-operative.

b. The maxims of conversation
Quality: Be truthful.

(i) Don't say what is false.
(ii) Don't say what lacks evidence.

Quantity: (i) Don't say less than is required.
(ii) Don't say more than is required.

Relation: Be relevant.
Manner: Be perspicuous.
(i) Avoid obscurity.
(ii) Avoid ambiguity.
(iii) Be brief.
(iv) Be orderly.

Assuming that the co-operative principle and its associated maxims are normally adhered to by both the speaker and addressee in a conversational interaction, Grice suggested that a conversational implicature—roughly, any meaning implied or expressed by and inferred or understood from the utterance of a sentence which is meant without being part of what is strictly said[1]—can arise from either strictly observing or ostentatiously flouting the maxims. In Huang (2003; 2007), I called conversational implicatures that are engendered by way of directly observing the maxims conversational implicatures$_O$, as in (2), and conversational implicatures that are generated by way of the speaker's deliberately flouting the maxims conversational implicatures$_F$, as in (3). (I use "+>" to stand for "(*ceteris paribus*) conversationally implicate".)

(2) John put on his hat and went out.
+>John first put on his hat and then went out

(3) John: Marshall is a bastard!
Mary: Oh, what a lovely day!
+> e.g., One shouldn't speak ill of people behind their back

A second Gricean dichotomy, independent of the conversational implicature$_O$/ conversational implicature$_F$ one, is between those conversational implicatures which arise without requiring any particular contextual conditions and those which do require such conditions. Grice called the first kind generalized conversational implicatures (GCIs), as in (4), and the second kind particularized conversational implicatures (PCIs), as in (5).

(4) The earthquake killed some of the villagers.
+> The earthquake did not kill all of the villagers

(5) John: Where's Peter?
Mary: The light in his office is on.
+> Peter is in his office

[1] Saul (2002) is of the view that Grice's main goal is to develop a theory of speaker-meaning. Following Saul, Horn (2004; 2006: 24) now holds that contra his own earlier work, conversational implicature is a component of speaker-meaning rather than a pragmatic inference (but see also Horn 2006: 35). By contrast, Levinson (2000), Atlas (2005), and others are still treating conversational implicature as a pragmatic inference. My definition is applicable to both sides.

The theoretical importance of this Gricean dichotomy has recently been subject to heated debates. Hirschberg (1991), Welker (1994) and Carston (2002), for example, doubted whether such a distinction can be maintained. On the other hand, Levinson (2000) put forward a spirited defense of it (see also Traugott 2004*a* for supporting evidence from semantic change). Finally, Grice designed a battery of tests to facilitate the identification of conversational implicatures. First, there is defeasibility or cancellability—conversational implicatures can simply disappear in certain linguistic or non-linguistic contexts. A second property exhibited by conversational implicatures is non-detachability—any linguistic expression with the same semantic content tends to carry the same conversational implicature. (A principled exception is those conversational implicatures that arise via the maxim of Manner, about which later.) Thirdly, calculability—conversational implicatures can transparently be derived via the co-operative principle and its attendant maxims. Fourthly, non-conventionality—conversational implicatures, though dependent on the saying of what is coded, are non-coded in nature. Fifthly, reinforceability—conversational implicatures can be made explicit without producing a sense of redundancy. Sixthly, some conversational implicatures may be indeterminate. They can be taken as conveying an open-ended range of implicatures relating to matters in hand. Finally, we have universality—conversational implicatures tend to be universal, being rationally motivated rather than arbitrary (see also Levinson 2000; Huang 2007: 32–5; Bach 2006*b*). In summary, Grice's account of conversational implicature is couched in a general theory of intention, co-operation, and more broadly, rationality.

24.2 The rise of neo-Gricean pragmatic theory of conversational implicature

While revolutionary in nature, what Grice presented at the James Williams Lectures was no more than just a sketchy proposal, albeit an ambitious one. As pointed out by Lakoff (1995: 194) metaphorically, "Grice himself provided an architect's sketch, but the full-fledged habitable edifice is still under construction; the original blueprint must be continually extended and reinterpreted to meet the needs of those who will actually inhabit it". Given Grice's seminal but sketchy proposal, it was no wonder that in the 1970s his ideas were considered by some scholars to be vague, superfluous, vacuous, unfounded, and even plain contradictory. Even Horn, himself a leading neo-Gricean, was of the following view:

> Grice's original framework is clearly at best incomplete and at worst inadequate beyond repair to the task of predicting sets of nonlogical inferences... It is simultaneously too

weak, in allowing the derivation of virtually anything by encompassing directly opposite maxims..., and too strong, in treating all calculable inferences monolithically.

(Horn 1988: 130)

Therefore, if the classical Gricean program was to be taken seriously within linguistics and the philosophy of language, much work had to be done to systematize, rigidify, and develop the original concepts that had been adumbrated by Grice. It was partially to meet this challenge that various neo-Gricean pragmatic reformations were developed.

What, then, have the neo-Griceans done to improve the classical Gricean pragmatic theory of conversational implicature? A number of areas can be identified. In the first place, individual types of classical Gricean conversational implicature were systematized. Horn (1972) represented the first attempt to provide a systematic analysis of conversational implicatures due to Grice's first sub-maxim of Quantity. He succeeded in providing a formalized account of what has since come to be known as $\text{Q-}_{\text{scalar}}$ implicatures, as in (4) above. The next major breakthrough relating to the same Gricean sub-maxim came from Gazdar (1979). Inspired in part by Horn's treatment of $\text{Q-}_{\text{scalar}}$ implicatures, Gazdar showed how $\text{Q-}_{\text{clausal}}$ implicature as in (6) can be formalized in an equally elegant way.

(6) John believes that his street voted Labour in the last election.
+> John's street might or might not vote Labour in the last election—the speaker doesn't know which

Later, Atlas and Levinson (1981) noted that if we apply the reasoning behind $\text{Q-}_{\text{scalar}}$ and $\text{Q-}_{\text{clausal}}$ implicatures to the implicatures specified in (2) above, we will get the wrong results. This led them to present the first formal analysis of conversational implicatures arising from the second half of Grice's sub-maxim of Quantity by appeal to a novel principle of informativeness ("Read as much into an utterance as is consistent with what you know about the world") (see also Atlas 2005), hence the term I-implicatures.

Secondly, more complex mechanisms were devised. For example, the constraints on Horn-scales (to be elaborated below), proposed by Levinson, successfully rules out *<regret, know>, *<iff, if>, and *<(p because q), (p and q)> as forming a genius Horn-scale. In the same vein, the Levinsonian resolution schema (to be discussed below) makes correct predictions for which type of conversational implicature overrides which type of conversational implicature under what circumstances. Thirdly, the Gricean maxims of conversation were reinterpreted. Martinich (1984) and Vanderveken (2002) linked each of the maxims to speech acts. Green (1996) recast them in terms of agency and intentionality. More recently, there have also been various attempts to integrate the classical and neo-Gricean pragmatic theories of conversational implicature with other current linguistic theories such as decision theory (Merin 1999), game theory (Benz et al. 2006), and bidirectional Optimality theory (Blutner and Zeevat 2004). Finally, the whole Gricean mechanism of the co-operative principle and its constituent maxims was submitted to various attempts at reduction. Early such attempts include Harnish (1976), in which Grice's maxims of Quality and Quantity are collapsed into a single maxim. Kasher (1976) argued that the entire Gricean machinery can be seen as following some sort of "most effective, least effort" rationality principle. More recently, Welker (1994) has tried to posit a super pragmatic principle.[2] This reductionist approach is consistent with the spirit of a metatheoretical desideratum known as "Occam's razor" which dictates that entities are not to be multiplied beyond necessity. However, of all the reductionist models, the most influential are the bipartite Hornian and the tripartite Levinsonian neo-Gricean typologies of conversational implicature, to which I now turn.

24.2.1 The Hornian typology

Horn (1984; 2007) put forward a dualistic model. On Horn's view, all of Grice's maxims (except the maxim of Quality) can be replaced with two fundamental and counterpoising principles: the Q[uantity]- and R[elation]-principles.

(7) Horn's Q- and R-principles
  a. The Q-principle
    Make your contribution sufficient;
    Say as much as you can (given the R-principle).
  b. The R-principle
    Make your contribution necessary;
    Say no more than you must (given the Q-principle).

In terms of information structure, Horn's Q-principle, which collects Grice's first sub-maxim of Quantity and his first two sub-maxims of Manner, is a lower-bounding pragmatic principle which may be (and characteristically is) exploited to engender upper-bounding conversational implicatures: a speaker, in saying "… *p* …", conversationally implicates that (for all he or she knows) "… at most *p* …". The *locus classicus* here is those conversational implicatures that arise from a prototype Horn-scale. Prototype Horn-scales are defined in (8) (Horn 1972; Atlas and Levinson 1981; Levinson 2000; Huang 2007: 38), with exemplification given in (9). Example (4) above is an instance of Q-implicatures.

(8) Prototype Horn-scales
  For $<S, W>$ to form a Horn-scale,

[2] Another influential reductionist model is, of course, Sperber and Wilson's (1995) unitarian relevance theory (RT). Given that RT does not endorse the general Gricean framework, and thus is not neo-Gricean, I do not discuss it in this chapter. See Huang (2007: 201–5) for a comparison between RT and the classical and neo-Gricean theory.

(i) A(*S*) entails A(*W*) for some arbitrary sentence frame A;
(ii) *S* and *W* are equally lexicalized, of the same word class, and from the same register;
(iii) *S* and *W* are "about" the same semantic relation, or from the same semantic field,

where *S* stands for "semantically strong expression" and *W* stands for "semantically weak expression".

(9) a. <all, most, many, some>
b. <hot, warm>
c. <beautiful, pretty, attractive>

On the other hand, the counterbalancing R-principle, which subsumes Grice's second sub-maxim of Quantity, his maxim of Relation, and his last two sub-maxims of Manner, and which is based on Atlas and Levinson's (1981) principle of informativeness, is an upper-bounding pragmatic law which may be (and systematically is) exploited to invite low-bounding conversational implicatures: a speaker, in saying "... *p*... ", conversationally implicates that (for all he or she knows) "... more than *p*...". This is illustrated in (2) above. However, more recently Horn (2007) has been of the view that the R-principle is not in itself subsumable under Grice's co-operative principle but under rationality.

Viewing the Q- and R-principles as a mere instantiation of Zipfian economy (Zipf 1949), Horn (1984; 2007) explicitly equated the Q-principle ("a hearer-oriented economy for the maximization of informational content") with Zipf's Auditor's Economy (the Force of Diversification, which tends toward a vocabulary of *m* different words with one distinct meaning for each word) and the R-principle ("a speaker-oriented economy for the minimization of linguistic form") with Zipf's Speaker's Economy (the Force of Unification, which tends toward a vocabulary of one word which will refer to all the *m* distinct meanings). The notion of Speaker's Economy is further distinguishable between mental inertia or paradigmatic economy (économie mémorielle) and articulatory/physical inertia or syntagmatic economy (économie discursive), hence internally dialectic in its operation. The former is concerned with the reduction in the inventory of mental lexicon; the latter with the reduction in the number of linguistic units (Martinet 1962: 139; 1964: 169; Horn 2007: 173–4). While the Auditor's Economy places a lower bound on the informational content of the message, the Speaker's Economy places an upper bound on its form. Furthermore, Horn argued, quoting Paul (1899), Martinet (1962; 1964), and Carroll and Tanenhaus (1975) as support, that the whole Gricean mechanism for pragmatically contributed meaning can be derived from the dialectic interaction (in the classical Hegelian sense) between the two mutually constraining mirror-image forces in the following way.

(10) Horn's division of pragmatic labour
The use of a marked (relatively complex and/or prolix) expression when a corresponding unmarked (simpler, less "effortful") alternate expression is available tends to be interpreted as conveying a marked message (one which the unmarked alternative would not or could not have conveyed).

In effect, what the communicative equilibrium in (10) basically says is this: the R-principle generally takes precedence until the use of a contrastive linguistic form induces a Q-implicature to the non-applicability of the pertinent R-implicature.

24.2.2 The Levinsonian typology

Horn's proposal to reduce Grice's maxims to the Q- and R-principles was called into question by Levinson (1987*b*, 1991, 2000). In Levinson's view, Horn failed to draw a distinction between what Levinson called semantic minimization ("Semantically general expressions are preferred to semantically specific ones") and expression minimization ("'Shorter' expressions are preferred to 'longer' ones").[3] Consequently, inconsistency arises with Horn's use of the Q- and R-principles. For example, in Horn's division of pragmatic labour, the Q-principle operates primarily in terms of units of speech production whereas elsewhere, in Horn-scales, for instance, it operates primarily in terms of semantic informativeness.

Considerations along these lines led Levinson to argue for a clear separation between pragmatic principles governing an utterance's surface form and pragmatic principles governing its informational content (but see Horn 2007 for a vigorous defense of his dualistic model and Traugott 2004*b* for supporting arguments).[4] He proposed that the original Gricean program (the maxim of Quality apart) be reduced to three neo-Gricean pragmatic principles: what he dubbed the Q[uantity]-, I[nformativeness]-, and M[anner]-principles. Each of the three principles has two sides: a speaker's maxim, which specifies what the principle enjoins the speaker to say and a recipient's corollary, which dictates what it allows the addressee to infer. Let me take them one by one.

(11) Levinson's Q-principle (simplified) (Levinson 2000; Huang 2007).
Speaker: Do not say less than is required (bearing the I-principle in mind).
Addressee: What is not said is not the case.

The basic idea of the Q-principle is that the use of an expression (especially a semantically weaker one) in a set of contrastive semantic alternates (such as a Horn-scale) Q-implicates the negation of the interpretation associated with the use of

[3] There is, of course, a strong tendency for the two distinct minimizations (or economies) to be conflated. This general correlation, in fact, follows directly from the Zipfian theory of economy. See, for example, Huang (1994; 2007: 40) for further discussion.

[4] Horn (2007: 179) argued that monist RT is implicitly dualistic in nature, given that relevance is measured in a minimax of give–take effort and effect.

another expression (especially a semantically stronger one) in the same set. Seen the other way round, from the absence of an informationally stronger expression, we infer that the interpretation associated with the use of that expression does not hold. Hence, the Q-principle is essentially negative in nature.

Three types of Q-implicature can then be identified: (i) $\text{Q-}_{\text{scalar}}$ implicatures, as in (4) above; (ii) $\text{Q-}_{\text{clausal}}$ implicatures, as in (6) above; and (iii) what I dubbed $\text{Q-}_{\text{alternate}}$ implicatures in Huang (2007). As mentioned above, $\text{Q-}_{\text{scalar}}$ implicatures are derived from Horn-scales. Next, $\text{Q-}_{\text{clausal}}$ implicatures are pragmatically enriched meanings of epistemic uncertainty. Like $\text{Q-}_{\text{scalar}}$ implicatures, $\text{Q-}_{\text{clausal}}$ implicatures also rest on a set of contrastive semantic alternates, but in this case, of a constructional kind. Finally, we have $\text{Q-}_{\text{alternate}}$ implicatures, which come from a non-entailment semantic (contrast) set. Roughly, we have two subtypes here. In the first, the lexical expressions in the set are informationally ranked, as in (12). Following Huang (2007), let me call Q-implicatures deriving from such a set $\text{Q-}_{\text{ordered alternate}}$ implicatures. By contrast, in the second subtype, the lexical expressions in the set are of equal semantic strength, as in (13). Let me term Q-implicatures thus induced $\text{Q-}_{\text{unordered alternate}}$ implicatures. Furthermore, Horn (2007: 168–70) distinguished two kinds of pragmatic strengthening: informative and rhetorical. While I-implicature (to be discussed below) increases both informative and rhetorical strength, Q-implicature is informatively but not rhetorically stronger than the sentence uttered without the implicature. This is evidenced by what Horn (2007) called rank orders.

(12) John tried to give up smoking.
+> John did not succeed in giving up smoking

(13) We teach French, German, and Russian here.
+> We don't, for example, teach Spanish here

Next, there is Levinson's I-principle.

(14) Levinson's I-principle (simplified) (Levinson 2000; Huang 2007).
Speaker: Do not say more than is required (bearing the Q-principle in mind).
Addressee: What is generally said is stereotypically and specifically exemplified.

Mirroring the effects of the Q-principle, the central tenet of the I-principle is that the use of a semantically general expression I-implicates a semantically specific interpretation. More accurately, the implicature engendered by the I-principle is one that accords best with the most stereotypical and explanatory expectation given our knowledge about the world.

(15) If you give me a free Beethoven, I'll buy five Mozarts.
+> If and only if you give me a free Beethoven will I buy five Mozarts

Finally, we come to Levinson's M-principle.

(16) Levinson's M-principle (simplified) (Levinson 2000; Huang 2007)
Speaker: Do not use a marked expression without reason.
Addressee: What is said in a marked way conveys a marked message.

Unlike the Q- and I-principles, which operate primarily in terms of semantic informativeness, the metalinguistic M-principle is operative primarily in terms of a set of alternates that contrast in form. The fundamental axiom upon which this principle rests is that the use of a marked expression M-implicates the negation of the interpretation associated with the use of an alternative, unmarked expression in the same set.

(17) a. The new manager is friendly.
I+> The new manager is friendly in the stereotypical sense

b. The new manager is not unfriendly.
M+> The new manager is less friendly than the utterance of (17a) suggests

Given the above tripartite classification of neo-Gricean pragmatic principles, the question that arises next is how inconsistencies arising from these potentially conflicting implicatures can be resolved. According to Levinson (2000), they can be resolved by an ordered set of precedence, which encapsulates in part the Hornian division of pragmatic labor.

(18) Levinson's resolution schema for the interaction of the Q-, I-, and M-principles
a. Level of genus: Q > M > I

b. Level of species: e.g., $Q\text{-}_{\text{clausal}} > Q\text{-}_{\text{scalar}}$

This is tantamount to saying that genuine Q-implicatures (where $Q\text{-}_{\text{clausal}}$ cancels rival $Q\text{-}_{\text{scalar}}$) supersedes inconsistent I-implicatures, but otherwise I-implicatures take precedence until the use of a marked linguistic expression triggers a complementary M-implicature to the negation of the applicability of the pertinent I-implicature. By way of summary, both Horn's and Levinson's neo-Gricean endeavors have put the classical Gricean theory on a much more rigorous basis, showing that the theory can be formalized and tested (or falsified), hence enhancing its predictive and explanatory adequacy.

24.3 Neo-Gricean pragmatics and lexicon

In the previous two sections, I outlined the classical and neo-Gricean pragmatic theories of conversational implicature. Starting from this section, I explore how neo-Gricean pragmatics can explain aspects of lexicon, semantics, and syntax. Let me begin with the lexicon.

24.3.1 Lexical narrowing

The meaning of a lexical expression can sometimes be narrowed down in the use of it, as can be illustrated by (19) and (20).

(19) a. John folded the newspaper into a rectangle.
b. +> John did not fold the newspaper into a square

(20) a. John had a glass of milk for breakfast this morning.
b. John had a glass of cow's milk for breakfast this morning

Lexical narrowing or strengthening can be grouped into two types. In the first, the use of the superordinate term of a hyponymic taxonomy where there is a specific hyponym denotes more narrowly the complement of the extension of the hyponym (Kempson 1980). This is the case for (19) (see (21)).

(21) rectangle

square rectangle

Lexical narrowing of this type follows directly from the Q-principle. Notice that *square* and *rectangle* form a Horn-scale here. Given the Q-principle, from the use of the semantically weaker *rectangle*, we obtain the pragmatically narrowed meaning "not square". This Q-based reduction of meaning typically gives rise to what Horn (1984) and Levinson (2000) called autohyponymy, i.e., privative polysemy. Other examples include *finger* +> "not thumb", *gay* +> "not lesbian", and *actor* +> "not actress". Note that these Q-narrowed meanings are not part of the lexical semantics of the items under consideration, because they can be canceled, as in (22). (I use "∼ +>" to signify "do not conversationally implicate".)

(22) John folded the newspaper into a rectangle, if not a square.
∼ +> John did not fold the newspaper into a square

Secondly, there is the I-implicature-based lexical narrowing. The basic idea here is that the use of a semantically general lexical item is I-implicated to a semantically more specific interpretation. This is the case for (20), where the semantically general term *milk* is I-narrowed to denote its culturally salient subset "cow's milk". Other examples include *secretary* +> "female secretary", *relationship* +> "sexual relationship", and *drink* +> "alcoholic drink". Of these, Horn (1984) and Levinson

(2000) were of the view that while *drink* is an autohyponym, *secretary* is not (see also Huang 1998; 2005).[5, 6]

24.3.2 Lexical blocking

Lexical blocking or pre-emption refers to the phenomenon whereby the appropriate use of a lexical expression formed by a relatively productive process is apparently prevented by the prior existence of a synonymous but distinct lexical item. Put the other way around, the existence of a conventional term for a particular meaning pre-empts or takes priority over any innovative expressions (Clark 1992). This process applies to both derivation and inflection. For example, *glory* partially blocks **gloriosity*, *hospitalize* (v) pre-empts **hospital* (v), and *went* fully blocks **goed*. Furthermore, lexical blocking can also take place between morphologically unrelated stems, as in *queen* precluding **kingess* and *thief* partially barring **stealer* (cf. *base stealer* in baseball).

Aronoff (1976) noted that the existence of a simple lexical expression can restrict the formulation of an otherwise expected affixally derived form with the identical meaning. This is the case for (23) and (24), where a preexisting simple abstract nominal underlying a given -*ous* pre-empts its nominalization with -*ity*.

(23) a. curious–curiosity
 b. furious–*furiosity (fury)

(24) a. tenacious–tenacity
 b. fallacious–*fallacity (fallacy)

Aronoff's analysis was, however, called into question by Kiparsky (1983) (see also Hofmann 1993, and Di Sciullo and Williams 1987 for a critique). On Kiparsky's view, Aronoff's account is both too strong and too weak. On the one hand, productive derivational processes are not always prevented by the existence of a more lexicalized alternative. This is evidenced by the fact that the abstract nominals *gloriousness* and *furiousness* co-exist peacefully with *glory* and *fury*. On the other hand, blocking is not limited to derivation but extends to inflection as well. As an alternative, Kiparsky suggested that Aronoff's blocking paradigm be reformulated as a lexical analog of the more general and ancient Elsewhere Condition, which can be traced at least back to Panini two millennia ago. However, the existence of partial blocking like *contestant/contester*, *informant/informer*, and *refrigerant/refrigerator*

[5] On Horn's (2007: 166) view, euphemism represents a *bona fide* case of culturally or socially motivated R-based narrowing.

[6] There is, of course, the other side of the lexical change coin, namely lexical broadening or loosening. According to Horn (2007: 165), this process of meaning expansion can be accounted for in terms of his R-principle. Within the relevance-theoretic framework, two types of lexical broadening are identified: approximation and category extension. See Wilson and Carston (2007) for an attempt to provide a unified account of lexical narrowing and broadening within the RT framework.

shows that Kiparsky's prediction ("Special rules block general rules in their shared domain") is still too powerful, because partial blocking corresponds to the phenomena that "the special affix occurs in some restricted meaning and the general affix picks up the remaining meaning" (Kiparsky 1983, see also Horn 1984). As an attempt to accommodate these cases, Kiparsky put forward a generalization which he dubbed "avoid synonymy" (see also Clark's 1990; 1992; 1993 principle of contrast).

(25) Kiparsky's avoid synonymy condition
The output of a lexical rule may not be synonymous with an existing lexical item.

What (25) basically predicts is the preemption of potential synonyms by established terms. As an initial illustrating example, consider *cook* and *cooker*. Given the established meaning of *cook* (a person who cooks), it constrains *cooker* with that meaning but not with a different meaning "an appliance which cooks". This represents partial blocking. Another case in point may involve the singular and plural forms of *fish*. The usual plural form of *fish*, namely *fish*, blocks *fishes* as its plural form, but does not block it when its sense is "different kinds of fish". Our final example comes from lexical change. After the Norman conquest in 1066, English speakers at court were faced with two sets of terms for animals: one from English (*calf*, *cattle*, *deer*, *pig*, *sheep* ...) and the other from French (*veau*, *boeuf*, *venaison*, *porc*, *mouton* ...). Given Kiparsky's avoid synonymy condition, the French terms were prevented from becoming synonymous with their English counterparts. Eventually, English and French terms are assigned different extensions: the original English terms for animals, but the terms borrowed from French for food. Thus we have the familiar pairs: *calf/veal*, *cattle/beef*, *deer/venison*, *pig/pork*, and *sheep/mutton* (see also Clark 1990). This has the effect that the use of a food-denoting term usually blocks the conceptual grinding mechanism with regard to the use of an animal-denoting one, as in (26).

(26) John doesn't like eating pork/?pig.

But lexical blocking of this kind can be canceled under certain conditions, resulting in what Blutner (2004) called deblocking. For example, Nunburg and Zaenen (1992) noted that the use of *cow* rather than *beef* is more appropriate in (27) (see also Copestake and Briscoe 1995 for further examples). This shows that there is essentially a pragmatic base for lexical blocking.

(27) Hindus are forbidden to eat cow/?beef.

We have a neo-Gricean pragmatic explanation for the complete and partial lexical blocking process ("Less productive/lexicalized/unmarked/irregular forms block more productive/lexicalized/marked/regular forms in the same slot") we have discussed so far. By Horn's division of pragmatic labour or Levinson's resolution schema, it is predicted that given the R/I-principle, unmarked forms tend to be used to convey unmarked messages, and given the Q/M-principle, marked forms to convey marked messages.

Next, in a pioneering study of the role played by Gricean conversational implicature in lexicon, which is independent of the work done by Aronoff and Kiparsky, McCawley (1978) discussed a number of cases of partial blocking, outside the area of derivation and inflection. One case concerns the formulation of color terms in English. As observed originally by Household, *pale red* is far less frequently used than, say, *pale blue*, *pale green*, and *pale yellow*. This is because while English has no lexical item for *pale blue*, *pale green*, and *pale yellow*, it has a lexical item for pale red, namely *pink*. Furthermore, what is interesting is the fact that *pale red* is found to be used occasionally. When it is used, it denotes a color other than pink, that is, a color that is paler than red but not as pale as pink (see also Huang 1998). This indicates that *pale red* is partially blocked by *pink*.

(28) a. John's girlfriend likes wearing pink skirts.
b. John's girlfriend likes wearing pale red skirts.
c. +> John's girlfriend likes wearing skirts whose color can't be described exactly as pink.

Again, the contrast shown between (28a) and (28b) falls out naturally from Horn's division of pragmatic labor or Levinson's resolution schema. While the use of (28a) engenders a straightforward I-implicated stereotypical interpretation, given that *pink* and *pale red* form an M-contrast set, the use of the marked (28b) M-implicates (28c). On the other hand, since there is no color term to block the use of, or to form an M-contrast set with *pale blue*, the use of *pale blue* (which means whitish blue) does not carry any M-implicated extra meaning.

Finally, following the analysis made by Shibatani of causatives in Japanese, McCawley pointed out that the distribution of a productive or periphrastic causative is also affected by the existence of a corresponding lexical causative. Whereas the use of a lexical causative, as in (29a), tends to depict a stereotypical, direct causative situation via the I-principle, the use of a productive causative, as in (29b), tends to refer to a more marked, indirect causation via the M-principle, hence the M-implicature in (29c) (see Haiman 1985 for the similar pattern in a number of languages other than English). As the reader can verify him- or herself, the complementary distribution between productive and lexical causatives is a direct reflex of the interaction between the I- and M-principles.

(29) a. John stopped the car.
b. John caused the car to stop.
c. +> John stopped the car in an unusual way.

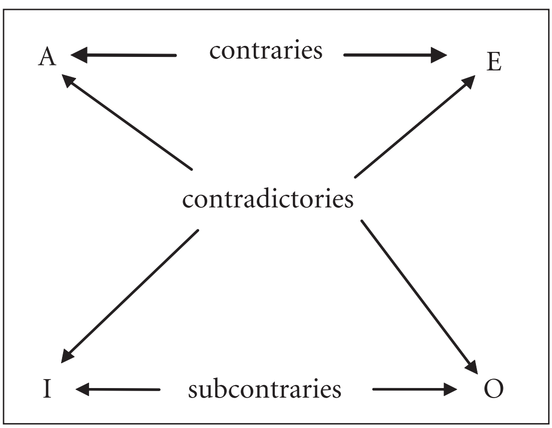
Again, as observed by McCawley, when there is no unmarked lexical causative, the M-implicated contrast does not appear. Consequently, the productive causative has a stereotypical interpretation, as in (30) (Levinson 2000: 141).[7, 8]

(30) a. *John laughen Mary.
b. John made Mary laugh.

24.3.3 Lexicalization asymmetry: Logical operators

Consider the traditional Square of Opposition formulated by Boethius out of Aristotle two millennia ago.

(31) Square of Opposition



Clearly, there is an asymmetry in the lexicalization of logical operators on the square. Whereas the A, E, and I vertices can all be lexicalized, the O vertex cannot. This is schematized in (32).

(32)

	A	I	E	O
Quantifiers	all	some	none	not all/*nall
Adverbs	always	sometimes	never	not always/*nalways
Connectives	and	or	nor	and not/*nand[9]
Modals	must	may	must not	permit not/*permitn't

Furthermore, as pointed out by Horn (1989: 252–67), the lexicalization asymmetry seems cross-linguistic. In other words, there is a strong cross-linguistic tendency

[7] Cf. *What had happened saddened/pleased John/made John (feel) sad/happy*. See also Horn (1984).

[8] As with the parallel case of pre-emption by synonymy, there is also pre-emption by homonymy. For instance, whereas we can say *They summered in Scotland*, we cannot say **They falled in Canada*. This is because the latter is blocked by the established, salient verb form *fall*, as used in *Something is falling from the sky* (Clark 1992). According to Horn (2007: 175), the tendency to avoid homonyms is Q-based.

[9] "NAND" or "nand" is the lexicalization of "and not" in a programming language.

for the O corner not to be represented in the lexicon. Rather, it is nearly always encoded by complex phrases.[10]

Why is this the case? To answer this question, we need to consider the question of what is the relation between the sub-contrary I and O vertices of the square? Put slightly differently, the question boils down to whether the I/O relation is a logical one or not. On Aristotle's view, the relation is logical as far as the modals are concerned, but, in the case of the quantifiers, it is non-logical. This was disputed by Hamilton and Jespersen, who held that the relation is logical for all the squares. By contrast, for De Morgan and J. S. Miller, it represents a non-logical relationship for all the squares (Levinson 2000: 68).

Now, given the neo-Gricean pragmatic theory, the answer to the puzzle presents itself. Notice that the vertices always form the scale of <A, I> and <E, O>. This has the consequence that the assertion of I Q-implicates "∼A", which is the contradictory of A, and hence equivalent to O. Using quantifiers, this can be illustrated in (33) and (34).

(33) <all, some>
Some monkeys have tails. (I)
+> Not all monkeys have tails (∼ A = O)

(34) <none, not all>
Not all monkeys have tails. (O)
+> Not none (i.e., some) monkeys have tails (∼ E = I)

Thus, the I and O corners are related by nothing but a generalized Q-$_{\text{scalar}}$ implicature. Furthermore, given that conversational implicatures are cancelable, the I/O relation is a non-logical one.

Let me now return to the question raised at the beginning of this sub-section, namely, why only the O value resists lexicalization? The answer is straightforward within the neo-Gricean pragmatic framework: what is Q-implicated on the square is not lexicalized. But such an explanation raises a further question, namely, given that the I and O vertices have the same communicational load, why the I corner can be and indeed is lexicalized. The answer, according to both Horn (2006) and Levinson (2000: 70–1), can be sought in the relatively complex nature or functional markedness of negation. Given a choice between a positive term and a negative term with the same communicational load, the positive term is usually picked up as the basic form to be lexicalized.

[10] This applies to natural languages only. See Horn (2006) for lexicalized O values such as NAND and XOR in non-natural languages.

24.4 Neo-Gricean pragmatics and semantics

In the last section, I looked at the role played by neo-Gricean pragmatics in the lexicon. In this section, I move to semantics.

24.4.1 $Q_{\text{-scalar}}$ implicatures: Pragmatic or semantic/ grammatical?

It has been noticed at least since Aristotle that a sentence like (35) has two systematically distinct interpretations: a one-sided, lower-bounded reading, as in (35a) and a two-sided, upper- and lower-bounded reading, as in (35b).

(35) The tea is warm.
   a. The tea is at least warm
   b. The tea is warm but not hot

How can a semanticist deal with sentences like (35)? He or she has to treat these sentences as lexically or logically ambiguous. However, there is a serious problem at the very heart of this ambiguity analysis, namely that it runs directly against the spirit of "Occam's razor". This has the consequence that, all things being equal, an account which has to suggest two lexical items is to be rejected in favor of an analysis which does not.

This is where neo-Gricean pragmatics comes in. As proposed in Horn (1972) and formalized in Gazdar (1979), the alternative analysis, which Horn (2006) dubbed the Golden Age of Pure Pragmatics (GAPP), is to obtain the one-sided, lower-bounded interpretation from semantics, but to derive the one-sided, upper-bounded reading via $Q_{\text{-scalar}}$ implicature. In other words, on this account, a sentence like (35) asserts or entails its one-sided, lower-bounded reading, Q-implicates its one-sided, upper-bounded reading, and the conjunction of the assertion and the implicature results in the corresponding two-sided, upper- and lower-bounded communicated understanding. This analysis applies to both logical operators and "ordinary" scalar predicates.[11]

But recently, this GAPP-style analysis has been challenged by Chierchia, Crain and their associates (e.g., Chierchia 2004; Crain and Pietroski 2002). On Chierchia and Crain's view, while a standard upper bounding $Q_{\text{-scalar}}$ implicature does arise from positive Horn-scales, it is quite weak and even blocked in negative Horn-scales

[11] In GAPP, cardinals are treated as scalar expressions by Horn (1972; 1989), but this GAPP analysis is now given up by him (Horn 2006; 2007). See, for example, Atlas (2005) Levinson (2000), Carston (2002), and Bultinck (2005) for further discussion.

and downward entailing environments (see (36a)). On the basis of this claim, Chierchia and Crain argued that Q-$_{\text{scalar}}$ implicatures must be computed compositionally. Consequently, they fall under compositional semantics, hence part of grammar, or innate linguistic mechanism.

(36) a. <not some, not many, not most, not all>
b. The earthquake didn't kill many of the villagers.
c. +> The earthquake killed some of the villagers

In fact, the observation made by Chierchia and Crain is not entirely novel. The projection properties of Q-$_{\text{scalar}}$ implicatures have long been a concern of neo-Griceans. As early as in 1979, Gazdar claimed that Q-$_{\text{scalar}}$ implicatures are suspended by logical operators (and) in embedded contexts. But as pointed out by Hirschberg (1991), Gazdar's generalization prevents too many Q-$_{\text{scalar}}$ implicatures. Hirschberg's own view is that Q-$_{\text{scalar}}$ implicatures are barred only under overt negation. But according to Horn (2006), Hirschberg's approach will block too few Q-$_{\text{scalar}}$ implicatures. Horn (1989: 233–4) suggested that Q-$_{\text{scalar}}$ implicatures are prevented in downward entailing contexts. However, Horn (2006) argued—convincingly I think—that, contra Chierchia and Crain, Q-$_{\text{scalar}}$ implicatures arising from negative Horn-scales, as in (36c), are not less robust than those which are derived from their positive counterparts. Furthermore, he acknowledged that it was Levinson (2000: 82, 254–5) who provided the correct answer. The alleged blockage of Q-$_{\text{scalar}}$ implicatures is due to the fact that a Horn-scale is reversed under negation and other downward entailing operators, as in (36a), and consequently a different Q-$_{\text{scalar}}$ implicature is derived from the inverse scale. If this is the case, then Chierchia and Crain's argument may not be maintained (see also Sauerland 2004 for a neo-Gricean analysis, and Horn 2006 and Ariel 2004; 2006 for the debate on the nature of *most*).

24.4.2 From the Gricean GCI/PCI dichotomy through a theory of three levels of meaning to a theory of presumptive meaning

Recall Grice's GCI/PCI dichotomy, discussed earlier. Based on this Gricean insight, Levinson (2000) developed a theory of presumptive meaning through a theory of three levels of meaning. On a traditional, standard view, there are only two levels of meaning to a theory of communication: a level of sentence-type-meaning vs. a level of utterance-token-meaning. The study of the former figures in semantics, and the study of the latter belongs to pragmatics. But Levinson (2000: 22) argued that such a view "is surely inadequate, indeed potentially pernicious, because it underestimates the regularity, recurrence, and systematicity of many kinds of pragmatic inferences". He proposed to add a third level—utterance-type-meaning—to the two generally accepted levels of meaning. This third layer is the level of generalized, preferred, or default interpretation, which is not dependent upon direct computations about speaker intentions but rather upon expectations about how language is characteristically used. GCIs, argued Levinson, should be included on this layer, as these pragmatic inferences have an expectable, stable, and even conventional interpretation. In order to account for this kind of conversational implicature, as we have already seen, Levinson has isolated a set of three default inferential heuristics—the Q-, I-, and M-principles, which is associated with a set of three default utterance-type conversational implicatures. Stated in this way, a neo-Gricean pragmatic theory of conversational implicature, which is largely concerned with generalized rather than particularized conversational implicatures, is essentially a theory of utterance-type-meaning on a level intermediate between sentence-type-meaning on the one hand, and utterance-token-meaning on the other. In other words, it is a theory of presumptive meaning—pragmatic inference that is generalized, default, and presumed.

However, as pointed by Levinson (2000), this middle layer of utterance-type-meaning has constantly been subject to attempts to reduce it, on the one hand, to the upper layer of sentence-type-meaning, as in, for example, Discourse Representation Theory, and on the other hand, to the lower level of utterance-token-meaning, as in, for example, Sperber and Wilson's (1986) relevance theory. In my view (Huang 2003, 2004*b*, 2007), such reductionist efforts, though methodologically highly desirable given the Occamistic principle discussed above, cannot be successful. The reason they will fail is this: on the one hand, GCIs are defeasible, that is, they can be canceled in certain linguistic and/or non-linguistic contexts. This will make it difficult for them to be semanticized. On the other hand, other things being equal, a theory about types is better than a theory about tokens in that the former enjoys more predictive and explanatory power. Therefore any attempts to reduce GCIs to nonce or once-off inferences should be resisted. If these arguments are correct, a three-tiered theory of communication with a layer of default interpretation sitting mid-way is in principle preferred over a two-leveled one without such an intermediate layer.

24.4.3 Pragmatic intrusion into what is said, Grice's circle, and the pragmatics–semantics interface

On a classical Gricean account, a distinction is made between what is said and what is conversationally implicated. Simply put, what is said is in general taken to be (i) the conventional meaning of the sentence uttered with the exclusion of any conventional implicature, and (ii) the truth-conditional content of a sentence uttered. However, according to Grice (1989: 25) and Levinson (2000: 172–86), before we work out what is said, we have to (i) resolve reference, (ii) fix deixis, (iii) disambiguate expressions, (iv) unpack ellipsis, and (v) narrow generalities.

What is conversationally implicated is then defined in contrast to, and calculated on the basis of, what is said (and in the case of M-implicatures, together with how what is said is said). Stated in this way, what is said is supposed to provide input to what is conversationally implicated.

It turns out, however, that the determination of (i)–(v) involves pragmatically enriched meaning of some kind. Put another way, there is pragmatic intrusion of some sort, namely the intrusion of pragmatically inferred content into the conventional, truth-conditional content, involved in the working out of what Grice called what is said.

The question that arises next is what is the pragmatic intrusion under consideration? Roughly, two current positions can be identified. The first is that the pragmatic intrusion is of a special kind, which differs from conversational implicature. Within this camp, three lines of arguments are of particular interest. According to Sperber and Wilson (1993), the pragmatic inference is an explicature, which is a development of the linguistically given logical form of the sentence uttered. Secondly, there is the position taken by Recanati (1993; 2004*a*; 2004*b*) that it is the pragmatically enriched part of what is said. A third argument is due to Bach (2004*b*), in which he proposed a third category of communicative content, intermediate between what is said and what is implicated. Bach dubbed the vehicle of such a content "implicIture", because it is implicit in what is said. Furthermore, Recanati put forward two tests, i.e., the availability and scope principles to differentiate explicature/the pragmatically enriched said/impliciture from conversational implicature.

The second position is represented by Levinson (2000). On Levinson's view, pragmatic intrusion into what is said is neither an explicature, nor the pragmatically enriched said, nor an impliciture. Rather, it is the same beast as a neo-Gricean conversational implicature. The reason is twofold. First, pragmatic intrusion into what is said is engendered by the same Gricean pragmatic mechanism that yields a conversational implicature. Secondly, as I argued in Huang (2007), neither of Recanati's tests works. Therefore, currently there is no reliable test that can be employed to distinguish alleged explicature/the pragmatically enriched said/impliciture from conversational implicature on a principled basis.

Following in the footsteps of work by Cohen, Wilson, Atlas and Gazdar, Levinson (2000) argued, biting the bullet, that contrary to Grice, conversational implicatures can intrude upon truth-conditional content. He showed how neo-Gricean conversational implicatures are involved in the working out of what is said, focusing on the determination of indexicality and related phenomena. Furthermore, he argued that the classic Cohen–Wilson argument can also be extended into logical connective constructions such as conditionals (37), comparatives (38), disjunctions (39), and *because*-clauses (40).

(37) If his son gets married and has children, John will be happy.

(38) Brushing your teeth and going to bed is better than going to bed and brushing your teeth.

(39) Mary either got married and had children or had children and got married—I don't know which.

(40) Because some of her friends came to her wedding, Mary was unhappy.

These constructions are labeled "intrusive" constructions by Levinson. The reason is that in these constructions, "the truth conditions of the whole depend in part on the implicatures of the parts" (Levinson 2000: 198). The truth-conditional content of (37)–(39) is dependent crucially on the generalized I-implicature stemming from the use of *and* to "and then". On the other hand, the quantifier *some* in (40) has to be Q-implicated to "some but not all" (but see Horn 2004; 2006 for a dissenting view of this non-GAPP account; see also King and Stanley 2005). Thus, there is no avoiding the conclusion that the truth condition of the complex construction has to be calculated taking into account the implicature of its part.

If neo-Gricean conversational implicatures can intrude onto truth-conditional content, this gives rise to a problem known as Grice's circle, namely how what is conversationally implicated can be defined in contrast to, and calculated on the basis of what is said, given that what is said seems to both determine and to be determined by what is conversationally implicated (see also Huang 2001; 2004*b*; 2007). Levinson's solution is that we should reject the "received" view of the pragmatics–semantics interface, according to which the output of semantics provides input to pragmatics, which then maps literal meaning to speaker-meaning. Rather, we should allow neo-Gricean pragmatics to play a systematic role in "pre"-semantics, i.e., to help determine the truth-conditional content of the sentence uttered. As Levinson (2000: 242) told us: "There is every reason to try and reconstrue the interaction between semantics and pragmatics as the intimate interlocking of distinct processes, rather than, as traditionally, in terms of the output of one being the input to the other." Such a radical proposal amounts to saying that the whole architecture of the theory of meaning needs to be radically reconstructed.

24.5 Neo-Gricean pragmatics and syntax: The case of anaphora and binding

Finally, I turn to the pragmatics–syntax interface, concentrating on anaphora and binding.

Anaphora

Anaphora can be defined as a relation between two linguistic elements, in which the interpretation of one (called an anaphoric expression) is in some way determined by the interpretation of the other (called an antecedent). Linguistic expressions that can be employed as an anaphoric expression include gaps (or empty categories), pronouns, reflexives, proper names, and definite descriptions.

Chomsky's binding conditions

Within the Principles-and-Parameters theory and its minimalist descendant, Chomsky (1995) distinguished two types of abstract feature for NPs: anaphors and pronominals. An anaphor is a feature representation of an NP which must be referentially dependent and which must be bound within an appropriately-defined minimal syntactic domain; a pronominal is a feature representation of an NP which may be referentially dependent but which must be free within such a domain. Interpreting anaphors and pronominals as two independent binary features, Chomsky hypothesized that we ideally expect to find four types of NP in a language—both overt and non-overt.

(41) Chomsky's typology of NPs

	Overt	Empty
a. [+anaphor, −pronominal]	lexical anaphor	NP-trace
b. [−anaphor, +pronominal]	pronoun	*pro*
c. [+anaphor, +pronominal]	–	PRO
d. [−anaphor, −pronominal]	name	*wh*-trace/variable

Of the three types of overt NP listed in (41), anaphors, pronominals, and r[eferential]-expressions are subject to binding conditions A, B, and C respectively.

(42) Chomsky's binding conditions
  A. An anaphor is bound in a local domain.
  B. A pronominal is free in a local domain.
  C. An r-expression is free.

Binding is defined in configurational terms, appealing to purely structural concepts like c-command, government, and locality. The binding theory is supposed to account for the syntactic distribution of the three types of overt NP listed in (41). Consider, for example, (43) from English.

(43) a. Newton$_1$ admired himself$_1$.
  b. Newton$_1$ admired him$_2$.
  c. Newton$_1$ admired Newton$_2$.

In (43a), *himself* is an anaphor in the Chomskyan sense. As such, it falls under binding condition A, according to which it is bound to its local antecedent *Newton*. Next in (43b), *him*, being a pronominal, is subject to binding condition B. Given binding condition B, it cannot be bound in its local domain, and there is thus disjoint reference between it and *Newton*. Finally, in (43c), the second *Newton* is an r-expression. By binding condition C, it cannot be co-indexed with the first *Newton*. From examples like these, Chomsky concluded that the syntactic distribution of anaphors, pronominals, and r-expressions is accounted for by binding conditions A, B and C, respectively. However, when confronted with a wider range of languages other than English, these binding conditions run into serious difficulties (see, for example, Huang 1991; 1992; 1994; 1995; 1996; 2000*a*; 2004*a*; 2006*b*; 2007, and Levinson 1987*b*; 1991; 2000 for detailed discussion).

The revised neo-Gricean pragmatic theory of anaphora

As an alternative to various syntactic and semantic approaches, a neo-Gricean pragmatic theory of anaphora was developed by Levinson (1987*b*; 1991; 2000) and Huang (1991; 1994; 2000*a*; 2000*b*; 2004*a*; 2006*b*; 2007). The central idea underlying the theory is that the interpretation of certain patterns of anaphora can be made utilizing pragmatically enriched meaning such as conversational implicatures, dependent on the language user's knowledge of the range of options available in the grammar, and of the systematic use or avoidance of particular anaphoric expressions or structures on particular occasions.

Applying the Q-, I-, and M-principles to the domain of anaphora, we can derive a revised neo-Gricean pragmatic apparatus for the interpretation of various types of anaphoric expressions.

(44) Huang's (2000*a*) revised neo-Gricean pragmatic apparatus for anaphora (simplified)

(i) The use of an anaphoric expression x I-implicates a local coreferential interpretation, unless (ii) or (iii).

(ii) There is an anaphoric Q-scale $< x, y >$, in which case the use of y Q-implicates the complement of the I-implicature associated with the use of x in terms of reference.

(iii) There is an anaphoric M-scale $\{x, y\}$, in which case the use of y M-implicates the complement of the I-implicature associated with the use of x, in terms of either reference or expectedness.

Needless to say, any interpretation generated by (44) is subject to the general consistency constraints applicable to conversational implicatures. These constraints include world knowledge, contextual information, and semantic entailments.

Let me now return to Chomsky's binding conditions and see how they can be re-interpreted in pragmatic terms. On the neo-Gricean pragmatic account, Chomsky's binding conditions B and C need not to be laid at the doorstep of generative syntax and can be reduced to pragmatics. In somewhat simplified terms, this can be achieved in the following way. If binding condition A is taken to be either grammatically constructed (as in the English-type, syntactic languages) or pragmatically specified via the I-principle (as in the Chinese-type, pragmatic languages), then binding condition B can be pegged directly to the application of the Q-principle. Given a speaker's knowledge of grammar and the I-principle, a reflexive will be chosen if coreference is intended. This has the consequence that if the reflexive is not employed but a pronoun is used instead, a Q-implicature will arise, namely no coreference is intended. In other words, we have a Horn-scale <reflexive, pronoun> here such that the use of a semantically weaker pronoun Q-implicates that the more informative, coreferential interpretation associated with the use of the reflexive cannot be truthfully entertained, as in (43b). By the same reasoning, binding condition C can also be eliminated. Wherever a reflexive could occur, the use of a semantically weaker proper name Q-implicates the non-applicability of the more informative, coreferential interpretation associated with the use of the reflexive. This is exactly what has happened in (43c). Furthermore, the revised neo-Gricean pragmatic theory can provide an elegant account of many of the anaphoric patterns that have embarrassed a generative analysis such as the case where contra binding condition B, a pronoun is bound in its local domain. In the case of long-distance reflexivization, the concept of logophoricity is invoked to explain why such a marked anaphoric expression is used. By logophoricity is meant the phenomenon whereby the "point of view" of an internal protagonist of a sentence or discourse, as opposed to that of the current, external speaker, is being reported using some morphological and/or syntactic means. The expression "point of view" is employed here in a technical sense and is intended to encompass words, thoughts, knowledge, emotion, and perception (e.g., Huang 2000*a*; 2002; 2006; 2007). This use of long-distance reflexives is accountable in terms of the M-principle. Since the grammar allows the unmarked pronoun to be employed to encode coreference, the speaker will use it if such a reading is intended, as in the Icelandic example (45a). On the other hand, if the unmarked pronoun is not used, but the marked long-distance reflexive is employed instead, then an M-implicature will be licensed. The implicature is that not only coreference but logophoricity as well is intended by the speaker. This is the case of (45b).

(45) (Icelandic, cited in Huang 2000*a*: 227)

a. Jon veit að Maria elskar hann.
John knows-INDIC that Mary loves-INDIC him
'John$_1$ knows that Mary loves him$_1$.'

b. Jon segir að Maria elski sig.
John says-INDIC that Mary loves-SBJV self
'John$_1$ says that Mary loves self$_1$.'

(INDIC = indicative mood, SBJV = subjunctive mood)

Notice another correlation here. If relevant, the choice between pronouns on the one hand and logophoric long-distance reflexives on the other is correlated with that between indicative and subjunctive mood in the embedded clause. The use of a pronoun tends to go with that of indicative mood, as in (45a); the employment of a logophoric long-distance reflexive tends to go with subjunctive mood, as in (45b). This correlation is a reflection of a semantic/pragmatic choice made by the external speaker about the responsibility he or she assumes for the truthfulness of what he or she is reporting. If a pronoun and indicative mood are used, it is indicative that the speaker asserts that the report is true. On the other hand, if a logophoric long-distance reflexive and subjunctive mood are deployed, it shows that the speaker does not take responsibility for the truth of the report (see, for example, Huang 2000*a*; 2002; 2006; 2007 for further discussion).

Since its inception, the neo-Gricean pragmatic theory of anaphora has been the impetus to a substantial amount of research and has been applied to a wide range of languages as genetically unrelated and structurally diverse as Korean, Spanish, and Turkish. In Huang (2000*a*; 2007) and Levinson (2000), substantial cross-linguistic evidence was presented to show that the revised neo-Gricean pragmatic theory of anaphora is more adequate than both a syntactic and a semantic approach. This indicates that pragmatics and syntax are intimately interconnected, though they are distinct levels and modes of explanation in linguistic theory. Contrary to the popular but erroneous Chomskyan view that syntax is autonomous, pragmatics plays a crucial role in explaining many of the phenomena that are thought to be at the very heart of syntax. If this is the case, then a large portion of linguistic explanation which is currently sought in syntactic terms may need to be shifted to pragmatics, hence the interaction and division of labor between pragmatics and syntax. This interface and division of labor may be summarized in a Kantian apophthegm: pragmatics without syntax is empty; syntax without pragmatics is blind (Huang 1994: 259; 2000*a*: 213; 2007). In addition, the revised neo-Gricean pragmatic theory of anaphora has important theoretical implications for universals, innateness, and learnability.

This page intentionally left blank

CHAPTER 25

PROBABILISTIC LINGUISTICS

RENS BOD

25.1 Introduction

25.1.1 Categorical versus probabilistic linguistics

Modern linguistic theory has evolved along the lines of the principle of categoricity: knowledge of language is characterized by a categorical system of grammar. Numbers play no role. This idealization has been fruitful for some time, but it underestimates human language capacities (Bresnan and Hay 2007). There is a growing realization that linguistic phenomena at all levels of representation, from phonological and morphological alternations to syntactic well-formedness judgments, display properties of continua and show markedly gradient behavior (see, for example, Wasow 2002; Gries 2005; Aarts 2007). While the study of gradience has a long history in the generative tradition (Chomsky 1955; 1961), recent approaches such as the Minimalist Program (Chomksy 1995) do not explicitly allow for gradience as part of the grammar (see Crocker and Keller 2006). Probabilistic

Many of the ideas presented in this chapter emerged from joint work carried out during the last fifteen years with (in chronological order): Remko Scha, Khalil Sima'an, Ronald Kaplan, Menno van Zaanen, Boris Cormons, Andy Way, Jennifer Hay, Stefanie Jannedy, Willem Zuidema, Gideon Borensztajn, Dave Cochran, Stefan Frank, and others. I am very grateful to all of them. Of course, any remaining errors in this chapter are entirely my responsibility. This work was partly funded by the Netherlands Organization for Scientific Research (NWO).

linguistics, instead, aims to focus on this relatively unexplored gradient middle ground.

But why should we model gradience by probabilities rather than by ranked rules, fuzzy set theory, connectionism, or yet another approach? One of the strongest arguments in favor of using probabilities comes from the wealth of frequency effects that pervade gradience in language (Bybee and Hopper 2001; Ellis 2002; Jurafsky 2003). Frequent words and constructions are learned faster than infrequent ones (Goodman et al. 2008). Frequent combinations of phonemes, morphemes and structures are perceived as more grammatical, or well-formed, than infrequent combinations (Coleman and Pierrehumbert 1997; Manning 2003). We can best model these effects by making explicit the link between frequency and probability: probability theory not only provides tools to working with the frequency of events but also with the frequency of *combinations* of events. In computing the probability of a complex event, such as a syntactic structure, we may not observe the structure in a store of previous language data. Probability theory allows for computing the probability of a complex event by combining the probabilities of their *subparts*.

Probabilistic linguistics is not just about modeling gradient linguistic phenomena, it also makes a cognitive claim. Following Bod, Hay, and Jannedy (2003), Gahl and Garnsey (2004), Jaeger and Snider (2008) and others, probabilities are an inherent part of the human language system. Probabilistic linguistics proposes that the language processing system is set up in such a way that, whenever an instance of a linguistic structure is processed, it is seen as a piece of evidence that affects the structure's probability distribution (Jaeger and Snider 2008).

While many linguists agree that there is a need to integrate probabilities into linguistics, the question is: Where? The answer in this chapter, as well as in other reviews, is: Everywhere. Probabilities are relevant at all levels of representation, from phonetics and syntax to semantics and discourse. Probabilities are operative in acquisition, perception, production, language change, language variation, language universals, and more. All evidence points to a probabilistic language faculty.

25.1.2 What does it mean to enrich linguistics with statistics?

To dispel dogmatic slumbers it may be good to realize that the main business of probabilistic linguistics is *not* to collect frequencies of words, collocations, or transitional probabilities. There is still a misconception that probabilities can be recorded only over surface events (see Manning 2003 for a discussion). Instead, there is no barrier to calculating probabilities over hidden structure, such as phrase-structure trees, feature-structures, or predicate–argument structures. Probabilistic linguistics *enriches* linguistic theory with statistics by defining probabilities over complex linguistic entities, from phonological to semantic representations. Probabilistic linguistics does therefore not abandon all the progress made by linguistics thus far; on the contrary, it integrates this knowledge with a probabilistic perspective.

One of the earliest successes of a probabilistic enrichment of a grammatical formalism is the *Probabilistic Context-Free Grammar* or PCFG (Grenander 1967; Suppes 1970). A PCFG consists of the simplest possible juxtaposition of a context-free grammar and probability theory: each context-free rule is enriched with a probability of application, such that the probability of a successive application of rules resulting in derivation of a sentence is computed by the product of the probabilities of the rules involved. For a long time, PCFGs had a less than marginal status in linguistics, which was partly due to the focus on categorical approaches in generative linguistics but also to the lack of annotated linguistic corpora needed for learning rule probabilities. Only during the last fifteen years or so have PCFGs led to concrete progress in modeling gradient linguistic phenomena, such as garden path effects (Jurafsky 1996), ambiguity resolution (Klein and Manning 2003), acceptability judgments (Crocker and Keller 2006), and reading times (Levy 2008). This progress crucially depended on the availability of linguistically annotated data (see Abeillé 2003 for an overview).

Despite this success, the shortcomings of PCFGs are also well acknowledged: their productive units capture only local dependencies while most syntactic phenomena involve non-local dependencies (see Joshi 2004). Furthermore, PCFGs correspond to the class of context-free languages while natural languages are known to be beyond context-free (Huybregts 1984). Although PCFGs have been useful in accurately parsing Penn Treebank sentences (e.g., Collins 1999; Charniak 1997), their cognitive relevance is much disputed (e.g., Fong and Berwick 2008).

Yet, the approach of *"stochasticizing" a grammatical formalism by enriching its grammatical units with probabilities* has been applied in many other formalisms, such as Tree-Adjoining Grammar, Combinatory Categorial Grammar, Lexical-Functional Grammar, and Head-Driven Phrase-Structure Grammar (e.g., Chiang 2003; Hockenmaier and Steedman 2002; Riezler and Johnson 2001). However, these probabilistic enrichments implicitly assume that the units of grammar coincide with the units of production and comprehension. Proponents of Construction Grammar, Cognitive Grammar, and usage-based linguistics have long emphasized that larger and more complex units play a role in language production and perception, such as conventional phrases, constructions, and idiomatic expressions (e.g., Kay and Fillmore 1999; Langacker 1987*b*; Barlow and Kemmer 2000; Bybee 2006*a*). What is needed is to assign probabilities to larger units of production, to which we will come back in the following section.

Instead of enriching the units of a grammatical formalism with probabilities, it is also possible to focus on a specific gradient phenomenon, next single out the possible factors that determine that phenomenon, and finally combine these factors into a probability model such as *logistic regression*. Logistic regression models are functions of a set of factors that predict a binary outcome (Baayen 2006). These models have been increasingly employed to deal with gradience in language production, such as in genitive alternation, dative alternation, presence/absence of complementizer (see Roland et al. 2005; Bresnan et al. 2007; Jaeger and Snider 2008). A logistic regression model permits simultaneous evaluation of all the factors in a model and assesses the strength of each factor relative to others. For example, in modeling ditransitive alternation between New Zealand and American English (e.g., in choosing between "*You can't give cheques to people*" vs. "*You can't give people cheques*"), Bresnan and Hay (2007) come up with a number of linguistic factors that may influence this syntactic choice, ranging from syntactic complexity, animacy, discourse accessibility and pronominality to semantic class. They next feed these factors to a logistic regression model, which indicates that NZ English speakers are more sensitive to animacy. Bresnan et al. (2007) furthermore show that their statistical model can correctly predict 94% of the production choices of the dative sentences in the 3-million-word Switchboard collection.

The method of logistic regression is flexible enough that it can be used for modeling a wide variety of other gradient phenomena, from grammatical choices in children's productions to syntactic persistence. However, logistic models require a set of predefined factors to begin with, rather than that they learn these factors from previous language experiences. Moreover, as with PCFGs, logistic models may have difficulties with global dependencies and larger units in language production. There is thus an important question whether these low-level models can be subsumed by a more general learning model.

Despite the differences of the statistical models discussed here, there is also a common view that emerges from these models and that may be summarized as follows: *Knowledge of language is sensitive to distributions of previous language experiences. Whenever an expression is processed, it is seen as a piece of evidence that affects the probability distribution of language experiences. New expressions are constructed by probabilistically generalizing over previous expressions.*

25.2 How far can probabilistic linguistics be stretched?

An approach that takes the direct consequence of the view above is Data-Oriented Parsing or DOP (Bod 1992; 1998; Scha et al. 1999; Kaplan 1996; and others). This approach analyzes and produces new sentences by combining fragments from previously analyzed sentences stored in a "corpus". Fragments can be of arbitrary size, ranging from simple context-free rules to entire trees, thereby allowing for both productivity and idiomaticity. The frequencies of occurrence of the fragments are used to compute the distribution of most probable analyses for a sentence (in perception), or the distribution of most probable sentences given a meaning to be conveyed (in production).

By allowing for all fragments, DOP subsumes other models as special cases, such as the aforementioned PCFGs (e.g., by limiting the fragments of trees to the smallest ones), as well as probabilistic lexicalized grammars (Charniak 1997) and probabilistic history-based grammars (Black et al. 1993). Carroll and Weir (2000) show that there is a subsumption lattice where PCFGs are at the bottom and DOP at the top. Moreover, DOP models can be developed for other linguistic representations, such as for HPSG's feature structures (e.g., Neumann and Flickinger 2002), LFG's functional structures (e.g., Arnold and Linardaki 2007), or TAG's elementary trees (e.g., Hoogweg 2003). DOP thus proposes a general method for "stochasticizing" a grammatical formalism.

25.2.1 An illustration of a generalized probabilistic model for phrase-structure trees

What does such a general model, which takes all fragments from previous data and lets frequencies decide, look like? In this section, we will illustrate a DOP model for syntactic surface constituent trees, although we could just as well have illustrated it for phonological, morphological, or other kind of representations. Consider a corpus of only two sentences with their syntactic analyses given in Figure 25.1 (we leave out some categories to keep the example simple).

On the basis of this corpus, the (new) sentence *She saw the dress with the telescope* can for example be derived by combining two fragments from the corpus—which we shall call *fragment trees* or *subtrees*—as shown in Figure 25.2. Note that there is no explicit distinction between words and structure in the subtrees. The combination operation between subtrees will for our illustration be limited to *label substitution*

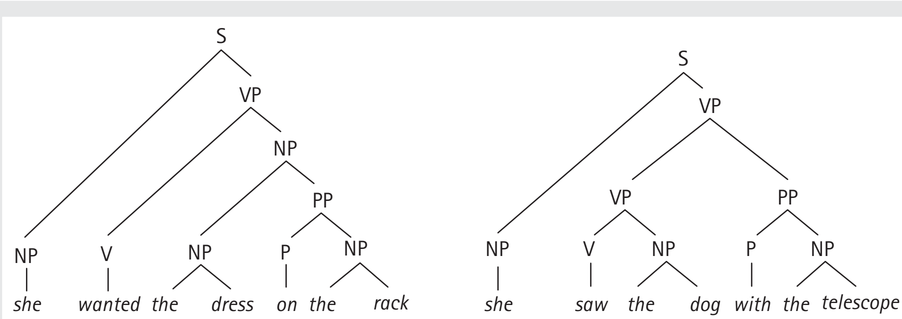


Figure 25.1. An extremely small corpus of two phrase-structure trees

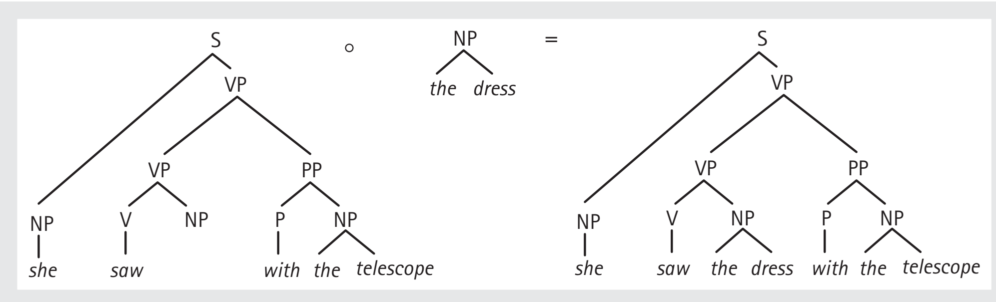


Figure 25.2. Analyzing a new sentence by combining subtrees from Figure 25.1

(but see below for extensions). This operation, indicated as $\circ$, identifies the leftmost nonterminal leaf node of the first subtree with the root node of the second subtree, i.e., the second subtree is substituted on the leftmost nonterminal leaf node of the first subtree provided that their categories match.

Thus in Figure 25.2, the sentence *She saw the dress with the telescope* is interpreted analogously to the corpus sentence *She saw the dog with the telescope*: both sentences receive the same phrase structure where the prepositional phrase *with the telescope* is attached to the VP *saw the dress*.

We can also derive an alternative phrase structure for the test sentence, namely by combining three (rather than two) subtrees from Figure 25.1, as shown in Figure 25.3. We will write $(t \circ u) \circ v$ as $t \circ u \circ v$ with the convention that $\circ$ is *left*-associative.

In Figure 25.3, the sentence *She saw the dress with the telescope* is analyzed in a different way where the PP *with the telescope* is attached to the NP *the dress*, corresponding to a different meaning from the tree in Figure 25.2. Thus the sentence is ambiguous in that it can be derived in (at least) two different ways, which is analogous either to the first tree or to the second tree in Figure 25.1.

Note that an unlimited number of sentences can be generated by combining subtrees from the corpus in Figure 25.1, such as *She saw the dress on the rack with the telescope* and *She saw the dress with the dog on the rack with the telescope*, etc. Thus

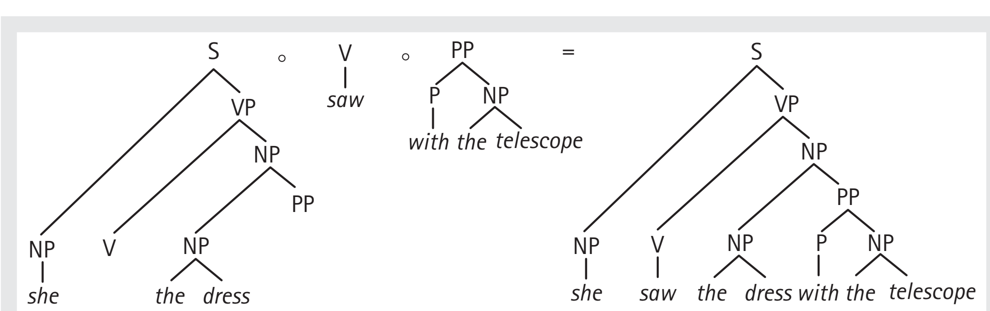


Figure 25.3. A different derivation for *She saw the dress with the telescope*

we obtain unlimited productivity by finite means. Note also that most sentences generated by DOP are highly ambiguous: many different analyses can be assigned to each sentence due to a combinatorial explosion of different prepositional-phrase attachments. Yet, most of the analyses are not plausible. They do not correspond to the interpretations humans perceive. Probabilistic linguistics proposes that it is the role of the probability model to select the most probable structure(s) for a certain utterance.

25.2.2 How to enrich a grammatical formalism with probabilities

How can we enrich the DOP model above with probabilities? By having defined a method for combining subtrees from a corpus of previous trees into new trees, we effectively established a way to view a corpus as a tree generation process. This process becomes a statistical process if we take the frequency distributions of the subtrees into account. For every tree and every sentence we can compute the probability that it is generated by this statistical process. Before we go into the details of this computation, let us illustrate the generation process by means of an even simpler corpus. Suppose that our example corpus consists of the two phrase-structure trees in Figure 25.4.

To compute the frequencies of the subtrees in this corpus, we need to define the (multi)set of subtrees that can be extracted from the corpus trees, which is given in Figure 25.5. Some subtrees occur twice in Figure 25.5: a subtree may be extracted from different trees and even several times from a single tree if the same node configuration appears at different positions. (Note that, except for the frontier nodes, each node in a subtree has the same daughter nodes as the corresponding node in the tree from which the subtree is extracted.)

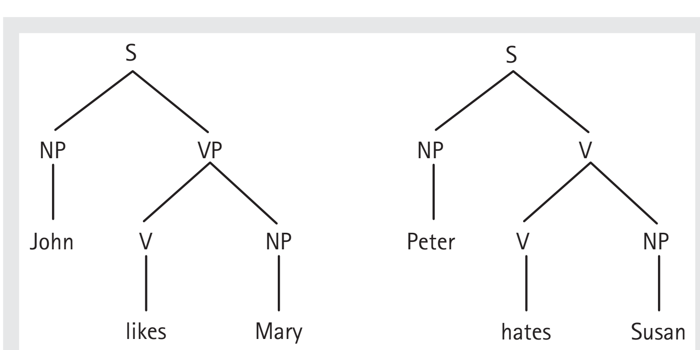


Figure 25.4. A corpus of two trees

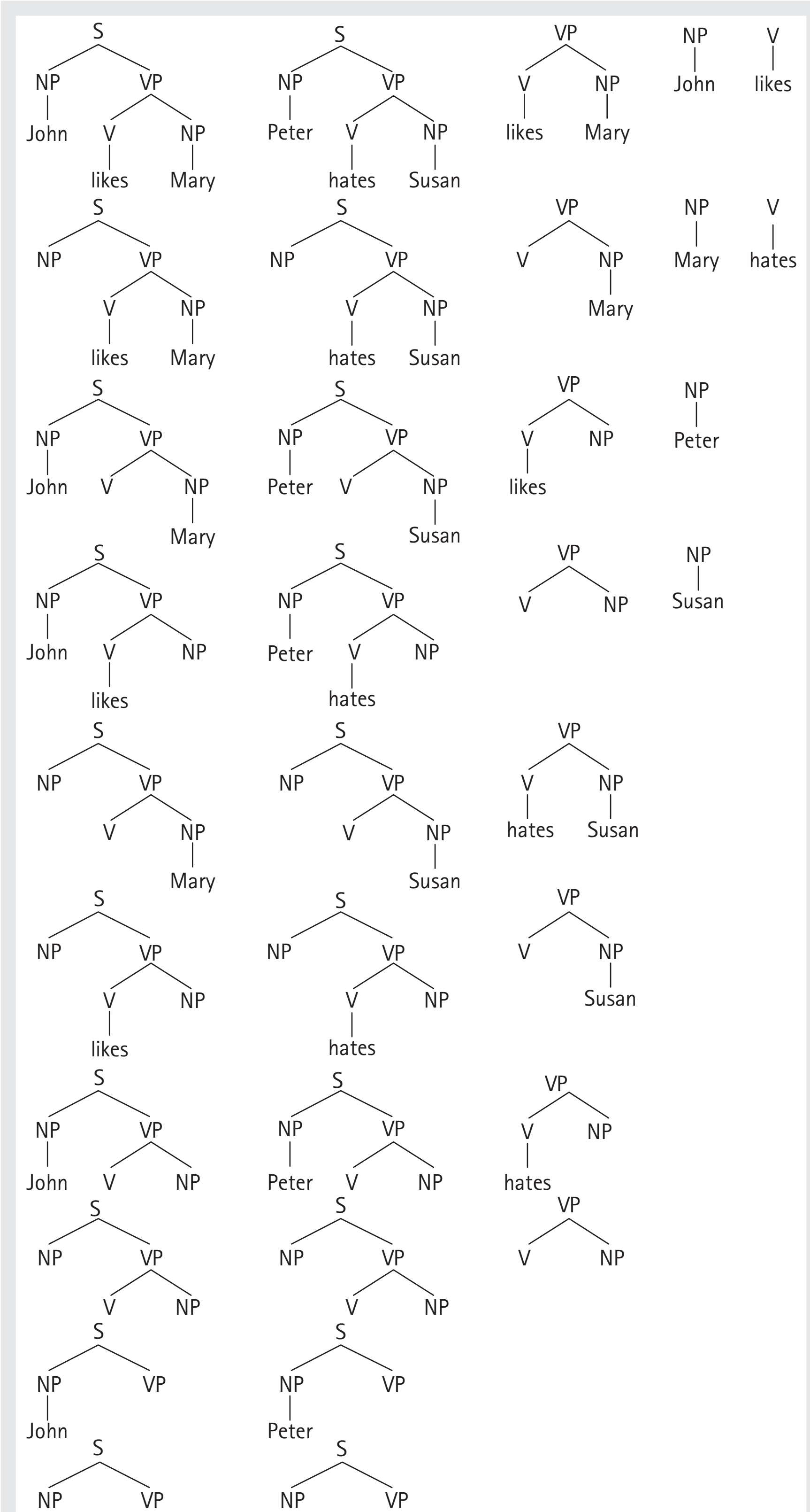


Figure 25.5. The bag of subtrees derived from the trees in Figure 25.4

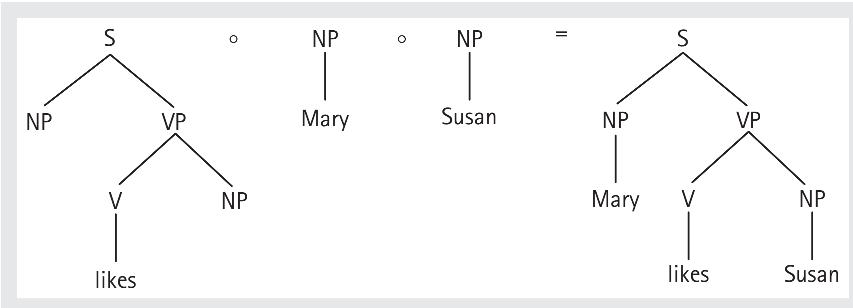


Figure 25.6. Analyzing *Mary likes Susan* by combining subtrees

As explained above, by using the substitution operation, new sentence-analyses can be constructed by means of this subtree collection. For instance, an analysis for the sentence *Mary likes Susan* can be generated by combining the three subtrees in Figure 25.6 from the set in Figure 25.5.

For the following it is important to distinguish between a *derivation* and an *analysis* of a sentence. By a derivation of a sentence we mean a sequence of subtrees the first of which is labeled with S and for which the iterative application of the substitution operation produces the particular sentence. By an analysis of a sentence we mean the resulting parse tree of a derivation of the sentence. Then the probability of the derivation in Figure 25.6 is the *joint* probability of three statistical events:

(1) selecting the subtree $_{\mathrm{S}}[\mathrm{NP}_{\mathrm{VP}}[_{\mathrm{V}}[\mathrm{likes}]\mathrm{NP}]]$ among the subtrees with root label S
(2) selecting the subtree $_{\mathrm{NP}}[\mathrm{Mary}]$ among the subtrees with root label NP
(3) selecting the subtree $_{\mathrm{NP}}[\mathrm{Susan}]$ among the subtrees with root label NP.

The probability of each event can be computed from the frequencies of the occurrences of the subtrees in the corpus. For instance, the probability of event (1) is computed by dividing the number of occurrences of the subtree $_{\mathrm{S}}[\mathrm{NP}_{\mathrm{VP}}[_{\mathrm{V}}[\mathrm{likes}]\ \mathrm{NP}]]$ by the total number of occurrences of subtrees with root label S: 1/20.

In general, let $|t|$ be the number of times subtree t occurs in the bag and $r(t)$ be the root node category of t, then the probability assigned to t is

$$P(t) = \frac{|t|}{\sum_{t':\, r(t')=r(t)} |t'|}$$

Since in our statistical generation process each subtree selection is independent of the previous selections, the probability of a derivation is the product of the probabilities of the subtrees it involves. Thus, the probability of the derivation in Figure 25.6 is: $1/20 \times 1/4 \times 1/4 = 1/320$. In general, the probability of a derivation $t_1 \circ \ldots \circ t_n$ is given by

$$P(t_1 \circ \ldots \circ t_n) = \prod_i P(t_i)$$

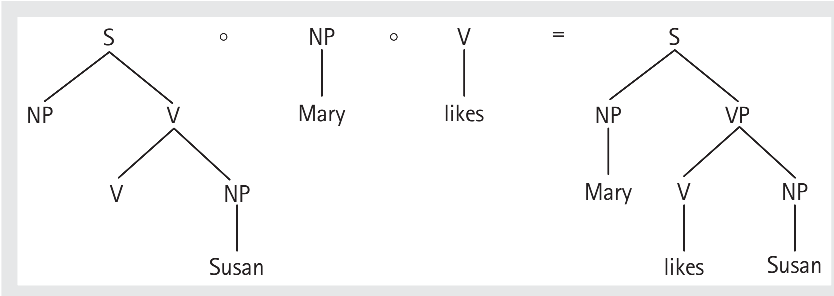


Figure 25.7. A different derivation yielding the same parse for *Mary likes Susan*

It should be stressed that the probability of an analysis or parse tree is *not* equal to the probability of a derivation producing it. There can be many different derivations resulting in the *same* parse tree. This "spurious ambiguity" may seem redundant from a linguistic point of view (and should not be confused with the "structural" ambiguity of a sentence). But from a statistical point of view, all derivations resulting in a certain parse tree contribute to the probability of that tree, such that no subtree that could possibly be of statistical interest is ignored.

For instance, the parse tree for *Mary likes Susan* derived in Figure 25.6 may also be derived as in Figure 25.7 or Figure 25.8. Thus, a parse tree can be generated by a large number of different derivations that involve different subtrees from the corpus. Each of these derivations has its own probability of being generated. For example, Table 25.1 shows the probabilities of the three example derivations given above.

The probability of a parse tree is the probability that it is produced by any of its derivations, also called the *disjoint* probability. That is, the probability of a parse tree T is the sum of the probabilities of its distinct derivations D:

$$P(T) = \sum_{D \text{ derives } T} P(D)$$

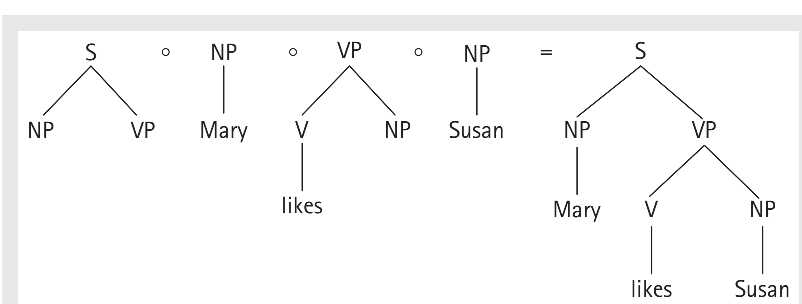


Figure 25.8. One more derivation yielding the same parse for *Mary likes Susan*

Table 25.1. Probabilities of the derivations in Figures 25.6, 25.7, and 25.8

P(Fig. 25.6)	$= 1/20 \times 1/4 \times 1/4$	$= 1/320$
P(Fig. 25.7)	$= 1/20 \times 1/4 \times 1/2$	$= 1/160$
P(Fig. 25.8)	$= 2/20 \times 1/4 \times 1/8 \times 1/4$	$= 1/1280$

Analogous to the probability of a parse tree, the probability of an utterance is the probability that it is yielded by any of its parse trees. This means that the probability of a word string W is the sum of the probabilities of its distinct parse trees T:

$$P(W) = \sum_{T \text{ yields } W} P(T)$$

For the task of language comprehension, we are often interested in finding the most probable parse tree given an utterance—or its most probable meaning if we use a corpus in which the trees are enriched with logical forms—and for the task of language production we are usually interested in the most probable utterance given a certain meaning or logical form. The probability of a parse tree T given that it yields a word string W is computed by dividing the probability of T by the sum of the probabilities of all parses that yield W (i.e., the probability of W):

$$P(T|T \text{ yields } W) = \frac{P(T)}{\sum_{T' \text{ yields } W} P(T')}$$

Since the sentence *Mary likes Susan* is unambiguous with respect to the corpus, the conditional probability of its parse tree is simply 1, by a vacuous application of the formula above. Of course a larger corpus might contain subtrees by which many different representations can be derived for a single sentence, and in that case the above formula for the conditional probability would provide a probabilistic ordering for them. For instance, suppose an example corpus contains the following trees given in Figure 25.9.

Two different parse trees can then be derived for the sentence *John hates buzzing bees*, given in Figure 25.10.

The DOP model will assign a lower probability to the tree 25.10 (a) since the sub-analysis 25.11 (a) of 25.10 (a) is not a corpus subtree and hence must be assembled from several smaller pieces (leading to a lower probability than when the sub-analysis was a corpus-subtree, since the probabilities of the pieces must be multiplied—remember that probabilities are numbers between 0 and 1). The sub-analysis 25.11 (b) of 25.10 (b) can also be assembled from smaller pieces, but it also appears as a corpus fragment. This means that 25.10 (b) has several more derivations than 25.10 (a), resulting in a higher total probability (as the probability of a tree is the sum of the probabilities of its derivations).

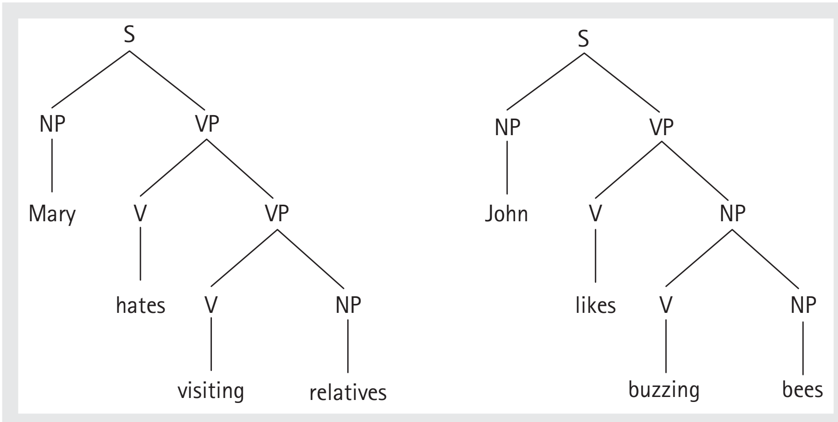


Figure 25.9. Two corpus trees for *Mary hates visiting relatives* and *John likes buzzing bees*

In general, there tends to be a preference in DOP for the parse tree that can be generated by the largest number of derivations. Since a parse tree which can (also) be generated by relatively large fragments has more derivations than a parse tree which can only be generated by relatively small fragments, there is also *a preference for the parse tree that can be constructed out of the largest possible corpus fragments*, and thus for the parse tree which is most similar to previously seen utterance-analyses (and note that the parse tree with the largest corpus subtrees also corresponds to the shortest derivation consisting of the *fewest* subtrees). The same kind of reasoning can be made for the probability of an utterance, i.e., there is a preference for the utterance (given a certain meaning or intention) that can be

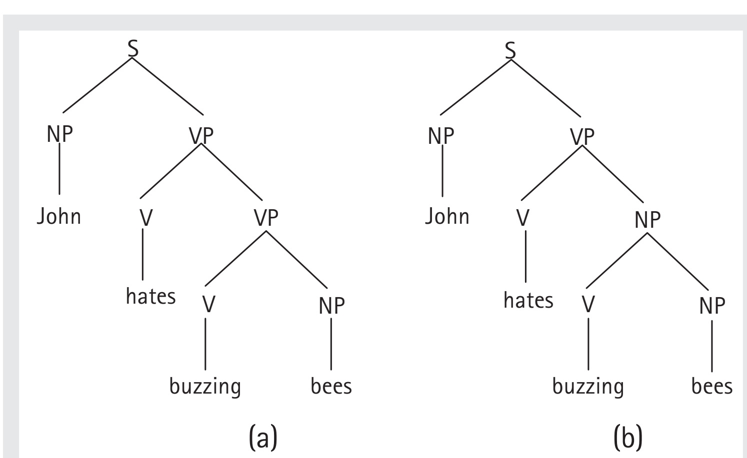


Figure 25.10. Parse trees for *John hates buzzing bees*

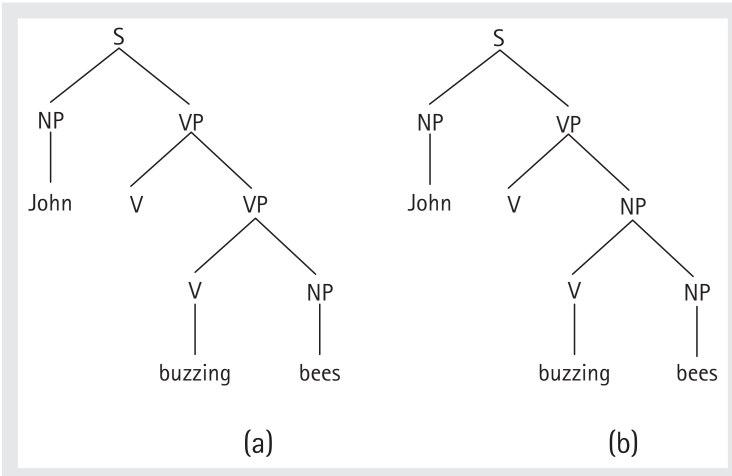


Figure 25.11. Two sub-analyses

constructed out of the largest possible corpus fragments, thus being most similar to previously seen utterances. This is particularly important to explain the use of constructions and prefabricated word combinations by natural language users (as we will discuss in section 25.3).

The notion of probability may be viewed as a measure for the *average similarity* between a sentence and the exemplars in the corpus: it correlates with the *number* of corpus trees that share fragments with the sentence, and also with the *size* of these shared fragments. DOP is thus congenial to analogical approaches to language that also interpret new input analogous to previous linguistic data, such as Skousen (1989) and Daelemans and van den Bosch (2005).

The probability model explained above is one of the simplest probability models for DOP, better known as "DOP1" (Bod 1998). DOP1 is "sound" in that the total probability mass of the sentences generated by the model is equal to one (Chi and Geman 1998; Bod, 2009). However, DOP1 has an inconsistent estimator (Johnson 2002): it can be shown that the most probable trees do not converge to the correct trees when the corpus grows to infinity. More advanced DOP models do have a consistent estimator such as Bod (2006*b*) or Zollmann and Sima'an (2005). Yet, these models still use DOP1 as a backbone; for example, the DOP model in Bod (2006*b*) starts with the subtree frequencies as in DOP1 that are next iteratively trained on a set of annotated sentences by the Expectation-Maximization algorithm (Dempster et al. 1977). It is important to stress that the definitions for computing the probability of a derivation, a parse tree, and a sentence are independent of the way the subtree probabilities are derived, and remain the same for different linguistic formalisms (see Bod, Scha and Sima'an 2003 for more details).

25.2.3 DOP models for richer grammatical formalisms

There is a common misconception that probabilities only deal with frequencies of events. On the contrary, probability models can incorporate many other factors, such as recency, meaning, and discourse context (Bod 1999), and, in Bayesian terms, probabilities can represent degrees of belief (e.g., Tenenbaum et al. 2006). Probability models have long been used in sociolinguistics (Labov 1966) and language change (see Bod, Hay, and Jannedy 2003), and they can also be defined over other grammatical frameworks, from Optimality Theory (Boersma and Hayes 2001) to Principles and Parameters theory (Yang 2004).

In this subsection, we will give a very short summary of a DOP model for a richer formalism just to show how such models can be developed in principle. In Bod and Kaplan (1998) we proposed a DOP model for the linguistically sophisticated representations used in LFG theory (Kaplan and Bresnan 1982). LFG representations consist of constituent structures and functional structures in correspondence. While constituent structures are labeled with simplex syntactic categories, functional structures also contain grammatical categories for subject, predicate and object, as well as agreement features and semantic forms, like predicate–argument structures. Figure 25.12 gives an example of a very simple corpus containing two

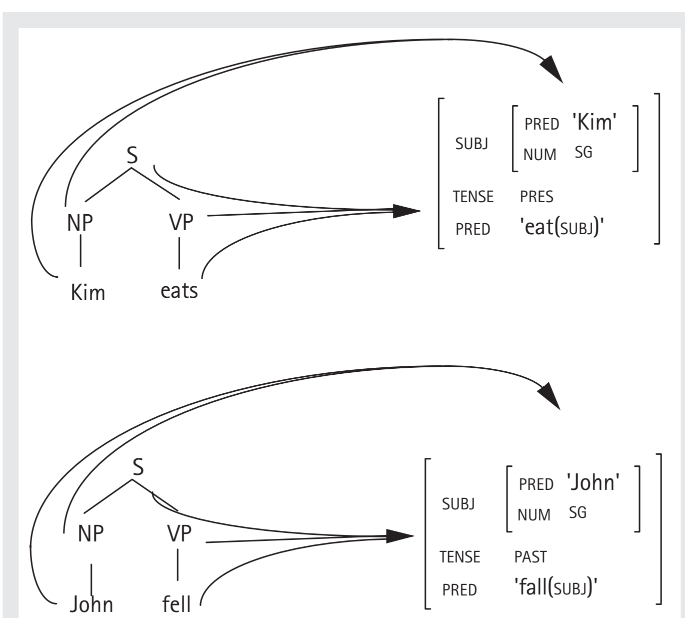


Figure 25.12. A corpus of LFG representations for *Kim eats* and *John fell*

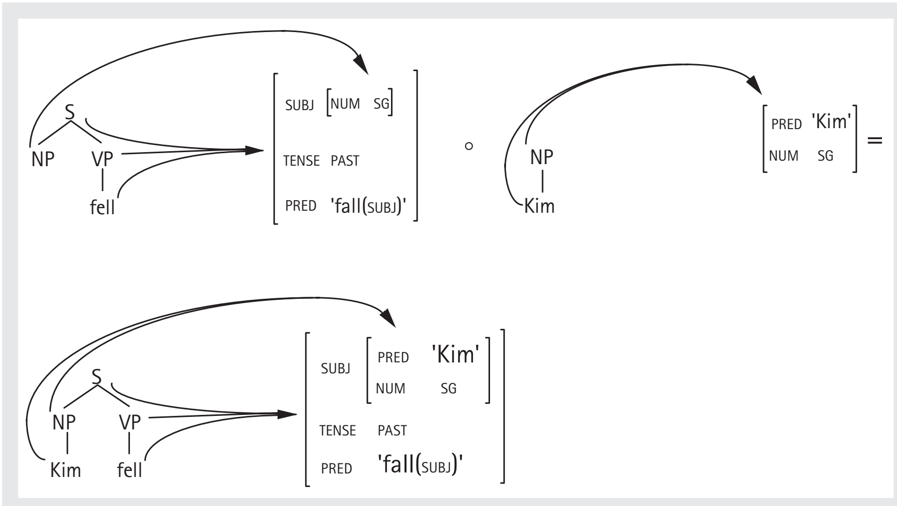


Figure 25.13. Deriving *Kim fell* in LFG-DOP from the corpus in Figure 25.12

LFG-representations for the sentences *Kim eats* and *John fell*, each of which consists of a constituent structure (a tree), a functional structure (an attribute-value matrix), and a mapping between the two (a correspondence function, represented by arrows).

As before, we take all fragments from these representations as possible productive units, and let statistics decide—but without using the derivational mechanism and rule-system provided by LFG theory. In this so-called "LFG-DOP" model, fragments are connected subtrees whose nodes are in correspondence with sub-units of f-structures (Bod and Kaplan 1998). For example, the following combination of fragments from the corpus in Figure 25.12 represents a derivation for the new sentence *Kim fell* (Figure 25.13).

The probability model for LFG-DOP can be developed along the same lines as in section 25.2.2, by assigning relative frequencies to the fragments and using the same definitions for the probability of a derivation and an analysis (see Hearne and Sima'an 2003 for more sophisticated fragment-estimation methods). Bod and Kaplan (1998) show how an interestingly different notion of "grammaticality with respect to a corpus" arises from LFG-DOP, resulting in a model which is both *robust*, in that it can parse and rank ungrammatical input, and which offers a formal account of meta-linguistic judgments such as *grammaticality* at the same time. Way (1999) and Arnold and Linardaki (2007) provide linguistic evaluations of LFG-DOP, while Finn et al. (2006) propose a computationally efficient approximation of LFG-DOP.

25.3 What can Data-Oriented Parsing explain?

25.3.1 Constructions and prefabs

DOP distinguishes itself from other probabilistic enrichments by taking into account constructions and units of arbitrary size. This allows the DOP approach to capture prefabs wherever they occur in the corpus (Manning and Schütze 1999: 446). For example, suppose that we want to produce a sentence corresponding to a meaning of asking someone's age (which in LFG-DOP is represented by the PRED value in the functional structure—see Figure 25.12). There may be several sentences with such a meaning, like *How old are you?*, *What age do you have?*, or even *How many years do you have?* Yet the first sentence is more acceptable than the other ones in that it corresponds to the conventional way of asking someone's age in English. This difference in acceptability is reflected by the different probabilities of these sentences in a representative corpus of English. While the probability of, for example, *What age do you have?* is likely to be small, since it will most likely not appear as a prefabricated unit in the corpus and has to be constructed out of smaller parts, the probability of *How old are you?* is likely to be high since it can *also* be constructed by one large unit. As we showed at the end of section 25.2.2, DOP's probability model prefers sentences that can be constructed out of the largest possible parts from the corpus. (And even in the case that both sentences should occur in a representative corpus of English, *How old are you?* would have the highest frequency.) Thus DOP prefers sentences and sentence-analyses that consist as much as possible of prefabs rather than "open choices".

25.3.2 Grammaticality judgments

In Bod (2001), DOP was tested against English native speakers who had to decide as quickly as possible whether three-word (subject–verb–object) sentences were grammatical. The test sentences were selected from the British National Corpus (BNC) and consisted of both frequent sentences such as *I like it* and low-frequency sentences such as *I keep it*, as well as sentences that were artificially constructed by substituting a word by another roughly equally frequent word, such as *I sleep it* and *I die it*, of which the grammaticality is dubious. Also a number of ungrammatical pseudo-sentences were added. It turned out that frequent sentences are recognized more easily and quickly than infrequent sentences, even after controlling for plausibility, word frequency, word complexity and syntactic structure. Next, an implementation of DOP was used to parse the test sentences. Each fragment *f* was assigned a response latency by its frequency *freq(f)* in the BNC as

$1/(1 + \log freq(f))$—see Baayen et al. (1997). The latency of the total sentence was estimated as the sum of the latencies of the fragments. The resulting model matched very well with the experimentally obtained reaction times (up to a constant) but only if *all* fragments were taken into account. The match significantly deteriorated if two-word and three-word chunks were deleted.

25.3.3 Disambiguation, interpretation, and translation

The experiments with grammaticality judgments described in section 25.3.2 trigger the hypothesis that "the accuracy of the model increases with increasing fragment size"—at least for grammaticality judgments. The hypothesis has now been corroborated also for modeling syntactic ambiguity (Bod 2001; Collins and Duffy 2001), translations from one language into another (Hearne and Way 2003), and the accuracy of semantic interpretation (Bod and Kaplan 2003). The hypothesis that the inclusion of larger productive units leads to better models has also been supported for languages other than English, that is, Dutch and French (Bod 1998; Cormons 1999), Hebrew (Sima'an et al. 2001), and Mandarin (Hearne and Way 2004). Furthermore, the hypothesis seems to be independent of the linguistic formalism: it was shown to be valid for LFG, HPSG, and TAG (see Bod, Scha and Sima'an 2003).

25.3.4 Syntactic priming and alternations

A possible challenge to DOP, and probabilistic linguistics in general, may seem to be the phenomenon of syntactic priming where it is the *low*-frequency rather than the high-frequency constructions that, when observed, have the highest chance of being primed. However, it should be kept in mind that the greatest change in a probability distribution is caused not by observing a high-frequent structure but by a low-frequent structure. Jaeger and Snider (2008) show that low-frequency constructions prime more as they result in a bigger change in the probability distribution, which in turn leads to an increased probability of reusing the same structure. Moreover, Snider (2008) develops a DOP model that integrates structural and lexical priming in language production. His model, coined DOP-LAST, is an extension of DOP1 with Exemplar Theory that can deal both with dative alternations and complex voice (active/passive) alternations.

25.3.5 Predicting the productive units

Although DOP starts from the assumption that any fragment can constitute a productive unit (and that large fragments are important), it can also make explicit predictions about the productive units that are actually used by humans in producing new sentences. Zuidema (2006) develops a DOP model that starts out with all subtrees, but that aims at finding the smallest set of productive units that explain the occurrences and co-occurences in a corpus. Large subtrees only receive non-zero weights if they occur more frequently than can be expected on the basis of the weights of smaller subtrees. In this way, Zuidema is able to make predictions about multi-word units and constructions used in adult language such as "I'd like to X", "from X to Y", "What's X doing?", etc. Borensztajn et al. (2008) test Zuidema's DOP model on child-produced utterances from the CHILDES database (MacWhinney 2000), where they split each corpus (for Eve, Sarah, and Adam) into three consecutive periods. It is found that the most likely productive units predicted by DOP closely correspond to the constructions found in empirical child-language studies by Tomasello (2003) and Lieven et al. (2003). In particular, Borensztajn et al. (2008) show that the DOP-derived productive units get more abstract with age (i.e., the number of open slots in the units increases across different periods). This corresponds to the empirical observation that children move from very concrete, item-based constructions ("holophrases") to more abstract contructions with open positions. We will come back to DOP and language acquisition in the next section.

It often happens that the productive units predicted by DOP look counter-intuitive, such as a subtree that is lexicalized only with the subject-noun and the determiner of the object with all other lexical elements as open slots. Yet it turns out that there are constructions where the subject has scope on the object's determiner, for instance in *She sneezed her way to the allergist* where the subject and the possessive determiner must be coreferential (**She sneezed his way to the allergist*) (Goldberg, p.c.).

25.4 How can probabilistic linguistics deal with language acquisition?

Probabilistic linguistics, as discussed so far, does not say anything about how the first structures are learned. It deals with statistical enrichments of linguistic formalisms on the basis of a corpus of already *given* structures. There is thus an important question of how we can extend our probabilistic approach to the problem of language acquisition.

Previous probabilistic learning models have been mostly based on a principle attributed to Harris (1954): word sequences surrounded by equal or similar contexts are likely to form the same constituent (e.g., van Zaanen 2000; Clark 2001; Klein and Manning 2005). While this idea has been fruitful, it has mostly been limited to *contiguous* contexts. For example, the Constituent-Context Model (CCM) by Klein and Manning (2005) is said to take into account "all contiguous subsequences of a sentence" in learning constituents (Klein and Manning 2005: 1,410). But this means that CCM neglects dependencies that are *non*-contiguous, such as between *closest* and *to* in "Show me the *closest* station *to* Union Square". Such non-contiguous dependencies are ubiquitous in natural language, ranging from particle verbs, agreement to auxiliary inversion.

There is a growing realization that non-linear, structural contexts must be included into a model of language learning (e.g., Culicover and Novak 2003; Dennis 2005; Solan et al. 2005; Seginer 2007). Below we will discuss how such contexts can be integrated in a general DOP framework for language learning.

25.4.1 A DOP model for language acquisition: U-DOP

We can extend DOP to language learning in a rather straightforward way, which is known as "Unsupervised DOP" or "U-DOP" (Bod 2006*b*, 2007*a*). If a language learner does not know which phrase-structure tree should be assigned to a sentence, he or she initially allows for all possible trees and lets linguistic experience decide which is the most likely one. As a first approximation we will limit the set of all possible trees to unlabeled binary trees. However, we can easily relax the binary restriction, and we will briefly come back to learning category labels in the next section. Conceptually, we can distinguish three learning phases under U-DOP:

(i) Assign all possible (unlabeled binary) trees to a set of given sentences
(ii) Divide the binary trees into all subtrees
(iii) Compute the most probable tree for each sentence.

The only prior knowledge assumed by U-DOP is the notion of tree and the concept of most probable tree. U-DOP thus inherits the rather agnostic approach of DOP. We do not constrain the units of learning beforehand but take all possible fragments and let the most probable tree decide.[1] We will discuss below how such an approach generalizes over other learning models, but we will first explain U-DOP in some detail by describing each of the learning phases above separately.

(i) Assign all unlabeled binary trees to a set of sentences
Suppose that a hypothetical language learner hears the two sentences *watch the dog* and *the dog barks*. How could the learner figure out the appropriate tree structures for these sentences? U-DOP conjectures that a learner does so by allowing (initially) any fragment of the heard sentences to form a productive unit and to try to reconstruct these sentences out of most probable combinations.

[1] In Bod (2009), we also propose a further extension of U-DOP which takes into account the shortest derivation as well.

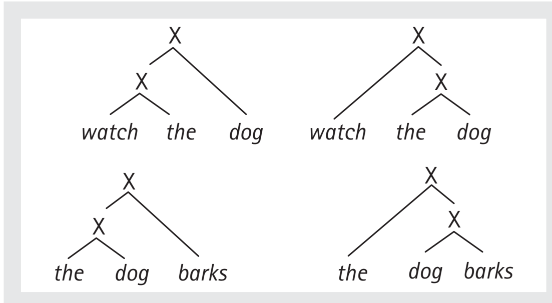


Figure 25.14. The unlabeled binary tree set for *watch the dog* and *the dog barks*

The set of all unlabeled binary trees for the sentences *watch the dog* and *the dog barks* is given in Figure 25.14, which for convenience we shall again refer to as the "corpus". Each node in each tree in the corpus is assigned the same category label X, since we do not (yet) know what label each phrase will receive. To keep our example simple, we do not assign labels to the words, but this can be done as well.

Although the number of possible binary trees for a sentence grows exponentially with sentence length, these binary trees can be efficiently represented in quadratic space by means of a "chart" or "shared parse forest", which is a standard technique in computational linguistics (see, for example, Kay 1980; Manning and Schütze 1999). However, for explaining the *conceptual* working of U-DOP, we will exhaustively enumerate all trees, keeping in mind that the trees are usually stored by a compact parse forest.

(ii) Divide the binary trees into all subtrees

Figure 25.15 lists the subtrees that can be extracted from the trees in Figure 25.14. The first subtree in each row represents the whole sentence as a chunk, while the second and the third are "proper" subtrees.

Note that while most subtrees occur once, the subtree $[the\ dog]_X$ occurs twice. The number of subtrees in a binary tree grows exponentially with sentence length, but there exists an efficient parsing algorithm that parses a sentence by means of all subtrees from a set of given trees. This algorithm converts a set of subtrees into a compact reduction which is linear in the number of tree nodes (Goodman 2003).

(iii) Compute the most probable tree for each sentence

From the subtrees in Figure 25.15, U-DOP can compute the most probable tree for the corpus sentences as well as for new sentences. Consider the corpus sentence *the dog barks*. On the basis of the subtrees in Figure 25.15, two phrase-structure trees can be generated by U-DOP for this sentence, shown in Figure 25.16. Both tree structures can be produced by two different derivations, either by trivially selecting

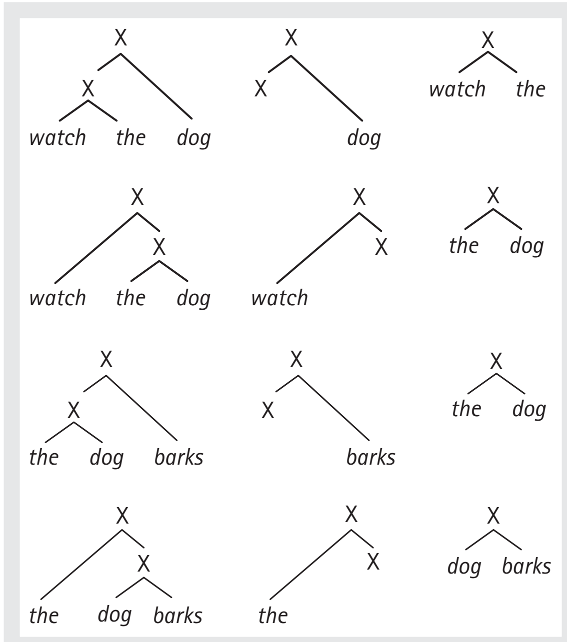


Figure 25.15. The subtree set for the binary trees in Figure 25.14

the largest possible subtrees from Figure 25.15 that span the whole sentence or by combining two smaller subtrees.

Thus the sentence *the dog barks* can be trivially parsed by any of its fully spanning trees, which is a direct consequence of U-DOP's property that subtrees of any size may play a role in language learning. This situation does not usually occur when structures for *new* sentences are learned.

U-DOP computes the most probable tree in the same way as the supervised version of DOP explained above. Since the subtree [*the dog*] is the only subtree that occurs more than once, we can informally predict that the most probable tree corresponds to the structure [[*the dog*] *barks*] where *the dog* is a constituent. This can also be shown formally by applying the probability definitions given in section 25.2. Thus the probability of the tree structure [*the* [*dog barks*]] is equal to the sum of the probabilities of its derivations in Figure 25.16. The probability of the first derivation consisting of the fully spanning tree is simply equal to the probability of selecting this tree from the space of all subtrees in Figure 25.15, which is $1/12$. The probability of the second derivation of [*the* [*dog barks*]] in Figure 25.16 is equal to the product of the probabilities of selecting the two subtrees which is $1/12 \times 1/12 = 1/144$. The total probability of the tree is the probability that it is generated by any of its derivations which is the sum of the probabilities of the

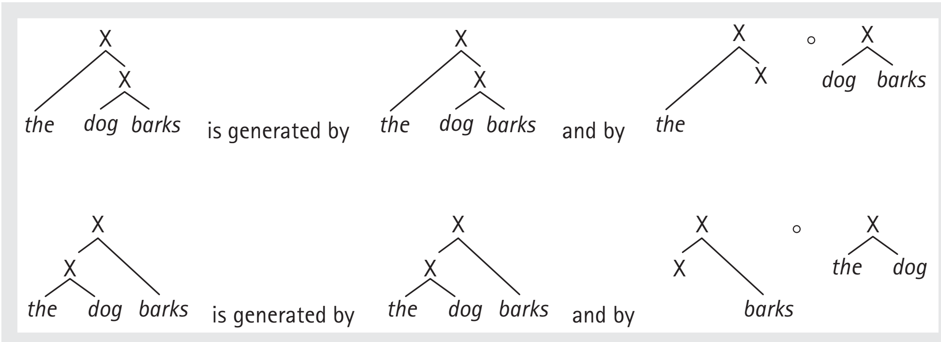


Figure 25.16. Parsing *the dog barks* from the subtrees in Figure 25.15

derivations:

$$P([the[dog\ barks]]) = 1/12 + (1/12 \times 1/12) = 13/144.$$

Similarly, we can compute the probability of the alternative tree structure [[*the dog*] *barks*], which follows from its derivations in Figure 25.16. Note that the only difference is the probability of the subtree [*the dog*] being 2/12 (as it occurs twice). The total probability of this tree structure is:

$$P([[the\ dog]\ barks]) = 1/12 + (1/12 \times 2/12) = 14/144.$$

Thus the second tree wins, although by just a little bit. We leave the computation of the conditional probabilities of each tree given the sentence *the dog barks* to the reader (these are computed as the probability of each tree divided by the sum of probabilities of all trees for *the dog barks*). The relative difference in probability is small because the derivation consisting of the entire tree takes a considerable part of the probability mass (1/12).

For the sake of simplicity, we only used trees without lexical categories in our illustration of U-DOP. But we can straightforwardly assign abstract labels *X* to the words as well. If we do so for the sentences in Figure 25.14, then one of the possible subtrees for the sentence *watch the dog* is given in Figure 25.17. This subtree has a discontiguous yield *watch X dog*, which we will therefore refer to as a *discontiguous subtree.*

Discontiguous subtrees are important for covering a range of linguistic constructions, as those given in italics in sentences (1)–(5):

(1) BA carried *more* people *than* cargo in 2005.
(2) *What's* this scratch *doing* on the table?
(3) Don't *take* him *by surprise.*
(4) Fraser *put* dollie nighty *on.*
(5) Most software *companies* in Vietnam *are* small-sized.

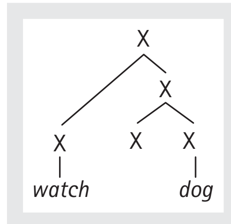


Figure 25.17. A discontiguous subtree

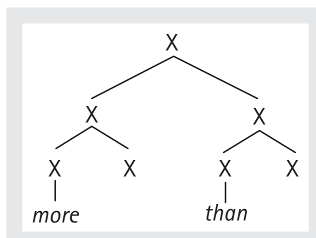


Figure 25.18. Discontiguous subtree for the construction *more . . . than . . .*

These constructions have been discussed at various places in the literature (e.g., Bod 1998; Goldberg 2006), and all of them are discontiguous. They range from idiomatic, multi-word units (e.g., (1)–(3)) and particle verbs (e.g., (4)) to regular syntactic agreement phenomena as in (5). The notion of subtree can easily capture the syntactic structure of these discontiguous constructions. For example, the construction *more* ... *than* ... in (1) may be represented by the subtree in Figure 25.18.

25.5 What can U-DOP learn?

25.5.1 Learning discontiguous phenomena

Discontiguous, structural dependencies play a role in virtually any facet of syntax. In Bod (2009), we show how U-DOP can learn particle verbs from child-directed speech in the Eve corpus (MacWhinney 2000), such as *blow* ... *up*, *take* ... *away*, *put* ... *on*, etc. For example, from the four child-directed sentences below, U-DOP can derive the dependency between *put* and *in*:

(1) *MOT: well we can put it in.

(2) *MOT: yeah.

(3) *MOT: Mom can put the stick in.

(4) *MOT: we just can't put any air in.

These sentences suffice for U-DOP to learn the construction *put X in*. At sentence 3, U-DOP induced that *can put it in* and *can put the stick in* are generalized by *can put X in*. At sentence 4, U-DOP additionally derived that *put X in* can occur separately from *can*, resulting in an additional constituent boundary. Thus by initially leaving open all possible structures, U-DOP incrementally rules out incorrect structures until the construction *put X in* is learned. Once the correct particle verb construction is derived, the production of incorrect constructions is blocked by the probability model's preference for reusing largest possible units given a meaning to be conveyed (assuming a semantic DOP model as in Bonnema et al. 1997).

25.5.2 Child language development from concrete to abstract

Note that in the examples above (but there are many more examples—see Bod, 2009), U-DOP follows a route from concrete constructions to more abstract constructions with open slots. These constructions initially correspond to "holophrases" after which they get more abstract resulting in the discontiguous phrasal verb. This is consonant with studies of child language acquisition (Peters 1983; Tomasello 2003) which indicate that children move from item-based constructions to constructions with open positions. The same development from concrete to abstract constructions has been quantitatively shown by Borensztajn et al. (2008) to hold for many other phenomena, including the use of the progressive, the use of auxiliaries, and do-support in questions and negations.

25.5.3 Learning rule-like behavior without rules: The case of auxiliary fronting

U-DOP can learn complex syntactic facets that are typically assumed to be governed by rules or constraints. Instead, in U-DOP/DOP, rule-like behavior can be a side effect of computing the most probable analysis. To show this we will discuss with some detail the phenomenon of auxiliary fronting. This phenomenon is often taken to support the well-known "Poverty of the Stimulus" argument and is called by Crain (1991) the "parade case of an innate constraint". Let's start with the usual examples which are the same as those used in Crain (1991), MacWhinney (2005), Clark and Eyraud (2006), and many others:

(5) The man is hungry.

If we turn sentence (5) into a (polar) interrogative, the auxiliary *is* is fronted, resulting in sentence (6).

(6) Is the man hungry?

A language learner might derive from these two sentences that the first occurring auxiliary is fronted. However, when the sentence also contains a relative clause with an auxiliary *is*, it should not be the first occurrence of *is* that is fronted but the one in the main clause:

(7) The man who is eating is hungry.

(8) Is the man who is eating hungry?

Many researchers have argued that there is no reason that children should favor the correct auxiliary fronting. Yet children do produce the correct sentences of the form (7) and rarely of the form (9) even if they have not heard the correct form before (Crain and Nakayama 1987).[2]

(9) *Is the man who eating is hungry?

According to the nativist view and the "poverty of the stimulus" argument, sentences of the type in (8) are so rare that children must have innately specified knowledge that allows them to learn this facet of language without ever having seen it (Crain and Nakayama 1987). On the other hand, it has been claimed that this type of sentence can be learned from experience alone (Lewis and Elman 2001; Reali and Christiansen 2005). We will not enter the controversy at this point (see Kam et al. 2005), but believe that both viewpoints overlook an alternative possibility, namely that auxiliary fronting needs neither be innately specified nor in the input data in order to be learned. Instead, the phenomenon may be a side effect of computing the most probable sentence-structure without learning any explicit rule or constraint for this phenomenon.

The learning of auxiliary fronting can proceed when we have induced tree structures for the following two sentences:

(10) The man who is eating is hungry.

(11) Is the boy hungry?

Note that these sentences do not contain an example of complex fronting where the auxiliary should be fronted from the main clause rather than from the relative clause. The tree structures for (10) and (11) can be derived from exactly the same sentences as in Clark and Eyraud (2006):

[2] Crain and Nakayama (1987) found that children never produced the incorrect form (9). But in a more detailed experiment on eliciting auxiliary fronting questions from children, Ambridge et al. (2008) found that the correct form was produced 26.7% of the time, the incorrect form in (9) was produced 4.55% of the time, and auxiliary doubling errors were produced 14.02% of the time. The other produced questions corresponded to shorter forms of the questions, unclassified errors, and other excluded responses.

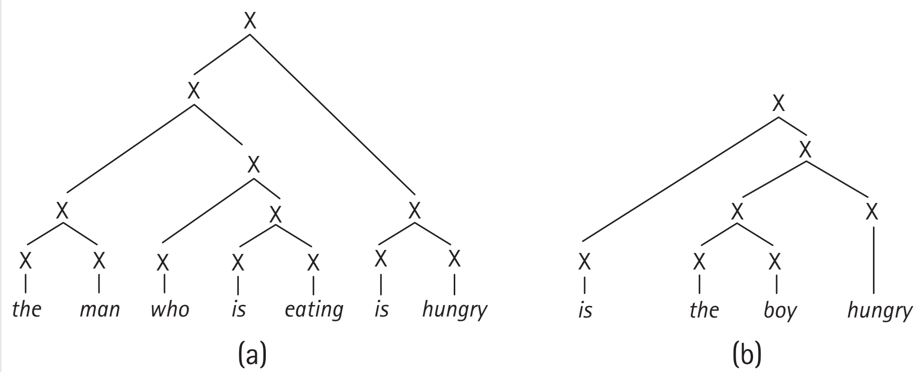


Figure 25.19. Tree structures for *the man who is eating is hungry* and *is the boy hungry?* learned by U-DOP from the sentences (10)–(15)

(12) The man who is eating mumbled.

(13) The man is hungry.

(14) The man mumbled.

(15) The boy is eating.

It can be shown that the most probable trees for (10) and (11) computed by U-DOP from sentences (10)–(15) are those in Figure 25.19 (see Bod, 2009, for details).

Given these trees, we can easily show that the most probable tree produces the correct auxiliary fronting. In order to produce the correct AUX-question, *Is the man who is eating hungry*, we only need to combine the following two subtrees in Figure 25.20 from the acquired structures in Figure 25.19 (note that the first subtree is discontiguous).[3]

Instead, to produce the incorrect AUX-question **Is the man who eating is hungry?* we would need to combine at least four subtrees from Figure 25.19, which are given in Figure 25.21.

The derivation in Figure 25.20 turns out to be the most likely one, thereby overruling the incorrect form produced in Figure 25.21. This may be intuitively understood as follows. We have already explained in section 25.4.2 that (U-)DOP's probability model has a very strong preference for sentences and structures that can be constructed out of largest corpus fragments. This means that sentences generated by a shorter derivation tend to be preferred over sentences that can only be generated by longer derivations.

We should keep in mind that the example above is limited to a couple of artificial sentences. It only shows that U-DOP/DOP can infer a complex auxiliary question

[3] We are implicitly assuming a DOP model which computes the most probable sentence given a certain meaning to be conveyed, such as in Bonnema et al. (1997) and Bod (1998).

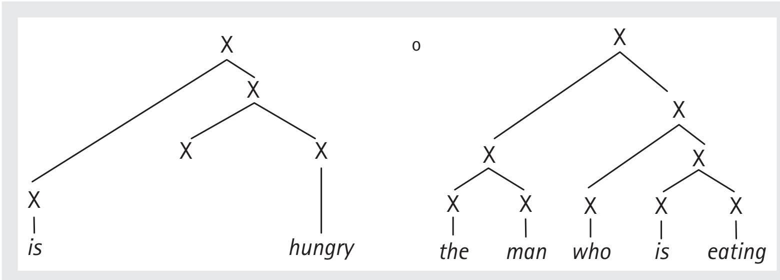


Figure 25.20. Producing the correct auxiliary fronting by combining two subtrees from Figure 25.19

from a simple auxiliary question and a complex declarative. But a language learner does not need to hear each time a new pair of sentences to produce a new auxiliary question—such as *Is the girl alone?* and *The girl who is crying is alone* in order to produce *Is the girl who is crying alone?* In Bod (2009), we show that U-DOP can also learn auxiliary fronting from the utterances in the Eve corpus (MacWhinney 2000), even though complex auxiliary fronting does not occur in that corpus. Furthermore, by sampling from the probability distribution of possible auxiliary sentences (rather than computing the most probable sentence as above), U-DOP can simulate many of the errors made by children as elicited in the experiments by Ambridge et al. (2008).

25.5.4 Learning categories and semantics?

Previous work has noted that category induction is an easier task than structure induction (Redington et al. 1998; Clark 2000; Klein and Manning 2005; Borensztajn and Zuidema 2007). The U-DOP approach can be generalized to category learning as follows. Assign initially all possible categories to every node in all possible trees (from a finite set of n abstract categories $C_1 \ldots C_n$) and let the most probable tree decide which trees correspond to the best category assignments (see Bod 2006*b*).

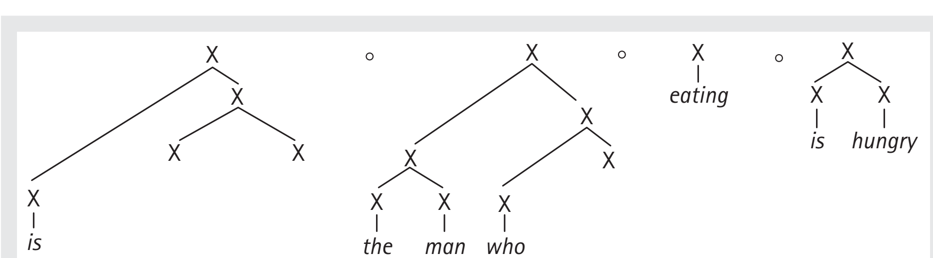


Figure 25.21. Producing the incorrect auxiliary fronting by combining four subtrees from Figure 25.19.

The unsupervised learning of semantic representations for sentences is still in its infancy. Although we can quite accurately learn meaning representations from a set of *pre-annotated* sentences (e.g., Bonnema et al. 1997; Bod 1999; Wong and Mooney 2007), the acquisition of semantic representations directly from child-directed speech is a largely unexplored field. Some progress has been made in (semi-)unsupervised learning of predicate–argument structure and compositional semantics (e.g., Alishahi and Stevenson 2008; Piantadosi et al. 2008). But it is fair to say that no satisfactory model for learning high-level semantic representations exists to date, neither in probabilistic nor in categorical linguistics.

25.5.5 Distinguishing possible from impossible languages?

There is an important question of whether probabilistic models, in particular U-DOP, don't learn too much. Can they learn impossible languages? Although this question has hardly been investigated so far, it is noteworthy that (U-)DOP's probability model, with its preference for the shortest derivation, limits the set of possible languages that can be learned. For example, a language that inverts a word string, called "linear inversion", will be ruled out by U-DOP. This is because linear inversion would lead to one of the longest derivations possible, since it can only be accomplished by decomposing the tree structures into the smallest subtrees for each single word, after which they must be reattached in the reverse order to the covering tree structure. Thus any structural operation leading to shorter derivations will win over this linear operation that tends to result in the longest possible derivation (at least for sentences longer than two words). While this property of U-DOP is promising, the relation between language typology and probabilistic learning models has still to be explored.

25.6 Relation to other models

As explained at the beginning of this chapter, probabilistic extensions can be created for virtually *any* linguistic theory or formalism. The underlying assumptions of (U-)DOP seem to be most congenial to Cognitive Grammar and Construction Grammar. DOP embraces the notion of "maximalist grammar" in cognitive and usage-based models, as coined by Langacker (1987*b*), and DOP is also consonant with Radical Construction Grammar where any exemplar or fragment is stored even if it is compositional (Croft 2001). However, we believe that both Cognitive Grammar and Construction Grammar suffer from a lack of formalization, especially in defining how constructions are combined and how they are learned—see Bod (2009) for a detailed criticism. At the same time, we have argued in Bod (2009) that DOP can be seen as a formalization and computational realization of Construction Grammar.

The link between DOP and exemplar/usage-based models is also straightforward. According to Exemplar Theory, stored linguistic tokens are the primitives of language that allow for production and perception as analogical generalizations over stored memories. Most exemplar models have been limited to phonetics (Johnson 1997; Bybee 2006*a*), but recent years have seen increasing interest in developing exemplar models for syntax, as evidenced by the papers in Gahl and Yu (2006). As Hay and Bresnan (2006) note, phonetic exemplar theory mainly deals with *classification* while syntactic exemplar theory (like DOP) focuses on *compositionality*. Schütze et al. (2007) aim to integrate the two approaches by extending similarity metrics from phonetic exemplar theory to syntax, which is congenial to the DOP-LAST model by Snider (2008).

Various linguists have argued that both rules and exemplars play a role in language, and have designed their linguistic theories accordingly (e.g., Sag and Wasow 1999; Jackendoff 2002; Culicover and Jackendoff 2005). The DOP model takes this idea one step further: It proposes that rules and exemplars are part of the same distribution and that both can be represented by fragment trees or subtrees. Rule-like behavior is then no more than a side effect of maximizing the probability from the frequencies of the subtrees. In language acquisition, U-DOP is most similar to Item-Based Learning (MacWhinney 1978; Tomasello 2003), especially in simulating the development from concrete item-based constructions to increasingly abstract constructions.

The first instantiation of DOP, DOP1 (see section 25.2), is formally equivalent to a Tree-Substitution Grammar or TSG. TSGs constitute a subclass of Tree-Adjoining Grammars or TAGs (Joshi 2004), and are equivalent to the class TAGs when DOP1's substitution operation is extended with adjunction (Hoogweg 2003). DOP can be seen as a TAG grammar where the elementary trees correspond to the set of *all* fragment trees derived from a treebank. One may wonder whether the learning method by Zuidema (2006)—where large subtrees only receive non-zero weights if they occur more frequently than can be expected from the weights of smaller subtrees (section 25.3.5)—turns the redundant DOP model into a linguistically succinct TAG model. But this is not the case. Zuidema's prediction of the productive units include redundant, overlapping fragments such as "I want X", "I'd like to X", "want from X to Y", "What's X doing?", etc. Without allowing for redundancy we cannot model gradient acceptability judgments (section 25.3.2) as these judgments are based on possibly overlapping constructions.

Probabilistic linguistics, with its emphasis on redundant, usage-based data, is of course in clear opposition to theories that emphasize a minimal set of non-redundant rules, in particular the Minimalist Program (Chomsky 1995). Yet, even there we may observe a converging trend. In their well-known paper, Hauser et al. (2002) claim that the core language faculty comprises just recursion and nothing else. If we take this idea seriously, then U-DOP may be the first computational model that instantiates it. U-DOP's trees encode the ultimate notion of recursion where every label can be recursively substituted for any other label. All else is statistics.

25.7 Conclusion

Probabilistic linguistics takes all linguistic evidence as positive evidence and lets statistics decide. It allows for accurate modeling of gradient phenomena in production and perception, and suggests that rule-like behavior is no more than a side effect of maximizing probability. Rules still appear in the scientific discourse but are not part of knowledge of language. According to this view, linguistic competence would consist not of a collection of succinctly represented generalizations that characterize a language; rather, competence may be nothing more than probabilistically organized memories of prior linguistic experiences.

We have seen that probabilistic models of language suggest that there is a single model for both language use and language acquisition. Yet these models need a definition of linguistic representation to start with, be it a phrase-structure tree or a functional attribute-value matrix. On this account, the central concern of linguistics would not be finding Universal Grammar but defining a *Universal Representation* for linguistic experiences that should apply to all languages. If there is anything innate in the human language faculty, it is this Universal Representation for linguistic experiences together with the capacity to take apart and recombine these experiences.

CHAPTER 26

LINGUISTIC RELATIVITY

ERIC PEDERSON

26.1 Introduction

This chapter presents not a linguistic theory or model but a body of research falling under the cover term *linguistic relativity*. Most generally, linguistic relativity studies investigate possible effects of natural language on purportedly non-linguistic cognition. For example, a linguistic relativity study might look to see whether speakers of a language which uses obligatory plural marking are more prone to remember numbers of objects in a visual display than speakers of a language which seldom marks plurality (cf. Lucy 1992*a*).

While linguistic relativity is defined here as a domain of research (much like morphology or syntax), much of the work in this area is rather ideological in tone and argument both for and against the possibility of an effect of language on cognition. Proponents of linguistic relativity generally express a faith that language effects on cognition are likely broad and important however difficult this may be to demonstrate. Opponents of linguistic relativity generally express great skepticism that language is likely to have much causal role in non-linguistic behavior and tend to seize on any specific failures to clearly demonstrate a specific language effect on cognition as indicative of a general lack of such effects. Both sides of the debate appeal to broad theoretical assumptions within linguistics, psychology, and cognitive science in general for a philosophical underpinning of their positions. As a result, those working to discover any linguistic relativity effects are not unlike scholars working within a particular and controversial model of language.

Just as collections of authors are often bundled together in a volume presenting work within a single theoretical model, so too, do we find books sharing a common interest in exploring linguistic relativity as an empirical question. See especially the collections by often allied researchers in Gentner and Goldin-Meadow (2003), Gumperz and Levinson (1996*a*), and Pütz and Verspoor (2000).

26.2 Historical background

Following the work of Edward Sapir, Benjamin Lee Whorf popularized the idea of linguistic relativity in the mid-20th century. From this, linguistic relativity is often termed the "Sapir–Whorf Hypothesis". This is somewhat of a misnomer for a number of reasons. One obvious problem is that many people have speculated on the effects of language on cognition in addition to Sapir and Whorf. Further, Whorf's writings on the topic were essentially written independently from Sapir, so the "Sapir–Whorf hypothesis" is not actually a joint statement put forward by these two. Furthermore, many popular conceptions of "Whorfianism" actually differ from Whorf's own speculations—see Smith (1996) for a discussion. More fundamentally, the Sapir–Whorf Hypothesis is not a genuine hypothesis. In science, a hypothesis is a specific claim formulated for empirical testing. One might have a hypothesis that native speakers of a particular language with a particular obligatory feature will be more sensitive to a corresponding feature in the environment or more likely to encode that feature in memory. On the other hand, a general statement that "language substantially influences thought" fails to be a falsifiable hypothesis.

There are nearly as many summaries of the history of linguistic relativity as there are articles describing original research in this area. For a general overview with particular sensitivity to cultural concerns see Hill and Mannheim (1992) and Lucy (1996; 1997). For a historical overview of Whorf's work and its relationship to modern work in linguistic relativity, see Lucy (1992*b*), Smith (1996), and Lee (1996; 2000). For a focus on the relationship of linguistic relativity to Cognitive Linguistics with particular concerns about methods, see Pederson (2007).

Historically, much of the debate about linguistic relativity has centered on a broadly imputed "strong vs. weak hypothesis". Certainly this is how the topic has been typically and cursorily treated in psychology textbooks. The alleged strong hypothesis states that language has a deterministic effect on the categories of cognition. Namely, the linguistic categories we learn as children lock us into congruent categories of thought. Given that people can successfully learn new second languages throughout their lives and indeed learn ways of thinking which were previously ineffable, the "strong" version is taken to be patently false. The "weak" version of the hypothesis states that there is an influence of linguistic categories in other areas of cognition. Work by Loftus and colleagues (Loftus and Palmer 1974; Loftus 1975) and more recently Lindquist and colleagues (Lindquist et al. 2006) studies the effects of immediate vocabulary choice on perception and memory within a single language. These studies amply demonstrate the power of language to influence non-linguistic cognition. Certainly good trial lawyers and authors know to manipulate lexical and constructional choice to their advantage. Such examples are taken to be true but trivial. That is, there is support for a "weak" version of linguistic relativity but this weak version is generally taken to be less interesting to cognitive science.

In short, the dichotomization of linguistic relativity concerns into a patently false strong position and a trivially true weak position only serves to reduce interest in the topic as a whole. Much of the work in modern linguistic relativity studies tries to avoid such broad oversimplification in favor of a more detailed model of language and cognition.

26.3 Requirements for linguistic relativity research

Clearly, it is no simple matter to determine what, if any effect, speakers' native languages have on their conceptualization of the world or on their cognitive patterns. Above all else, research in linguistic relativity requires considerable breadth of expertise (or co-operation among an interdisciplinary team of researchers). A description of the relevant features of at least two languages must be adequate to withstand the scrutiny of linguists working both with that language and in the domain under investigation (semantics/morphology/etc.). Further, it is not enough to observe that a language has a particular feature (e.g., obligatory plural marking or a morphological evidential system). The language description must also be sufficiently exhaustive to know to what extent certain concepts may be expressed and under what conditions (e.g., is there an optional plural marker and when is it used).

To this language expertise must be added adequate behavioral experimentation to determine patterns of cognition in different populations. This must be up to the standards of psychological research and may need to be conducted with population samples far from the usual laboratory setting. It can be quite challenging enough for psychologists to manage sufficient rigor in cross-cultural work; for anthropologists and linguists to satisfy a critical audience of psychologists may be especially difficult.

Contrasting with these challenging research parameters is the popular appeal of linguistic relativity. The lay public can be counted on to have an opinion on the topic (typically that language *does* affect patterns of thinking). This is scarcely surprising. To many non-linguists, it is self-evident that the purpose of language is to represent the world and that there are likely to be interesting variations in the ways in which languages do this. For example, bilingual speakers commonly report the subjective experience of "thinking differently" in their alternate languages. Unfortunately, it is perhaps impossible to evaluate precisely such statements.

In stark contrast to this common view, the modern fields of linguistics and psycholinguistics have been largely concerned with the purported universals of language. Further, most linguists scarcely concern themselves with semantics at all. The relatively few semanticists among them in turn typically eschew cross-linguistic comparison in favor of formalized descriptions relying critically on their own native speaker intuitions. The field of linguistics has long rewarded sophisticated theory development (most easily elaborated from work with better-known languages) far more than fieldwork-based description.

Semantics has long been one of the less empirical branches within linguistics. Linguistics has long provided excellent training for phonetics and structural analysis, but has a striking dearth of empirical methods for semantic description. So it is scarcely surprising that a typical descriptive grammar of a less-described language will give little attention to semantics. Semantic comparison across even moderately well-described languages largely relies on simple glossing conventions and dictionaries listing approximate translation equivalents. As a result, accurate cross-linguistic semantic descriptions are seldom available.

The first and most fundamental step in a linguistic relativity study must be to have an adequate description of the language categories of at least two appropriately distinct varieties of language. Failure to achieve this means that at best positive results will be open to multiple interpretations and at worst there will be no interpretable results at all from a failure to appropriately formulate a specific linguistic relativity hypothesis for testing. For instance, Loucks and Pederson (in press) argue that the lack of an adequate and appropriate linguistic description has precluded meaningful results in research on the categorization of motion events.

To create a linguistic description for linguistic relativity research, careful fieldwork with the languages must have been conducted. Existing semantic descriptions should generally be assumed inadequate or inaccurate. If one is interested in how habitual language patterns affect thought, then data about habitual language use must be collected rather than relying on grammatical treatises or dictionaries. For this, recordings of native speaker discussions concerning the domain in question should be collected and analyzed. Which aspects of what could be described are routinely selected for expression and which seem to be relatively ignored? Relying on extant recordings or transcripts is problematic. Since the goal is cross-linguistic comparison, it is best to rely on cross-linguistic data which derives from the same context repeated across speech communities. Early examples of such standardized cross-linguistic comparison can be found for child language in the “Frog story” paradigm (Berman and Slobin 1994) and much of this research ultimately led to discussions of motion event cognition (see section 26.7, below). For adult language the “Pear film” (Chafe 1980) has been widely used for elicitation of narratives in many language communities all speaking about the same sequence of events. Pederson et al. (1998) describe a method for eliciting and comparing very specific spatial systems across languages with the intended purpose of developing subsequent cognitive testing. While not immediately connected with cognitive testing, Bohnemeyer et al. (2007) use a similar research technique for developing a typology of motion event descriptions.

26.4 Linguistic relativity and mainstream linguistics

As linguistics and psychology joined forces in the cognitive sciences during the 1980s, there was a particular interest in determining what is universal to human cognition. Linguistics had already been greatly influenced by the “Chomskyan paradigm”, which sought to determine those native elements which are unique and necessary to any human language. Generally speaking, variation was theoretically interesting only insofar as it instantiated general universal principles. Over the years, the Chomskyan approach has been forced to reduce the number of features believed to be universal and unique to the human language faculty until the we reach the position in Hauser et al. (2002). In this controversial view, the only known feature remaining to uniquely structured language is an ill-defined “recursion” or a rule’s ability to refer to itself iteratively. For a passionate argument against the dominance of universalist/nativist approaches to language categories, see Levinson (2003*a*). For arguments specifically confronting Hauser et al. (2002), see Jackendoff and Pinker (2005*b*) and Bickerton (in press) among others.

Consistent with this universalist paradigm was the faith that language processing was essentially modular. That is, the processes of language production and comprehension were assumed to be essentially neurologically (that is, architecturally) autonomous from other cognitive processes—with the necessary exception of the input to and output from the language module. Neuroimaging and brain damage research during this period was often cited as supporting this view that language processing is largely autonomous from other brain processes. Thus, in the absence of clear evidence supporting linguistic relativity, it was simple enough to dismiss language—especially those aspects of language which are variable across languages—as having little influence on the rest of brain functioning.

There are a number of problems with such a view, however. Perhaps the most troubling is that the lack of empirical evidence for linguistic relativity effects may well have been simply due to a limited number of studies which had investigated the topic. As for the purported modularity of language—an autonomy from other brain processes—the picture even today is far less clear than one might hope. For all of their wonder, neuroimaging studies are still notably crude tools relative to the sophistication of their subject. Further, while there is some neuroanatomy which does indeed seem most architecturally dedicated to language processes, it is associated with fairly automatic processing of structural relationships, e.g., the processing of simple morphology and syntax. At its core, linguistic relativity concerns itself with the conceptual categories of language, that is to say, with the influence of habitual semantic processing. In contrast to simple grammatical parsing, semantic processing seems to involve neurology scattered over many regions of the cortex. This renders a strictly modular account of semantic processing less plausible.[1] Accordingly, it appears impossible to rule out interesting interactions between language processing and other conceptual processing on the basis of what is currently known about the human brain.

Arguments against linguistic relativity generally only argue for a deficiency in a "pro" linguistic relativity argument. Any such successful defeat of a single linguistic relativity study entails only that there is no evidence in this specific instance of a linguistic relativity "effect". One cannot generalize from a null result in one study to the conclusion that there can be no interesting effects in other linguistic and cognitive domains.

In other words, the irony of the "con" position is that linguistic relativity cannot be definitively disproven. One can always argue that the lack of results demonstrating a linguistic relativity hypothesis for a particular study should be taken as indicative that linguistic relativity effects would not be found elsewhere. However, such an argument remains largely an article of faith. Conversely, should a "pro" linguistic relativity study convince skeptics that an effect exists in a particular context and domain, it would be a similar article of faith to assume it exists elsewhere. Because of this, linguistic relativity researchers generally seek to find language effects in domains where skeptics would assume they would not be possible. To show an effect in a domain which is generally considered to be cross-culturally quite variable is less likely to impress than in an area in which cognitive universals have been presumed. Levinson and colleagues worked with spatial language

[1] There are countless, largely technical accounts of what is known of the neurology of language processes. The conclusions and the technologies are continually being updated, so the best first source for readers interested in this topic would be a current textbook in psycholinguistics.

and cognition for precisely this reason (see especially the discussion in Levinson (2003*b*; Levinson 1996*a*) as spatial language and cognition had been assumed to be largely invariant across languages and cultures. Research demonstrating variation in such a domain would naturally lead to serious consideration of a potentially large interplay between language and cognition. See the discussion in section 26.7 below.

There has been a recent and overdue trend toward formulating more precise hypotheses about specific conceptual categories in contextualized usage and what possible effects one's native language may have on the use and availability of these categories. Working with color terminology, Kay and Regier (2007) argue that a more complex model of the interactions between the cognitively universal and the linguistically specific must be developed. Similarly, Imai and Mazuka (2003) seek a more complex model for balancing the cognitively universal and language specificity in their work with the individuation of objects (as opposed to mass/substance). Arguably, the lack of such subtle models has been an obstacle to linguistic relativity research in the past. In other words, individual linguistic relativity studies should never expect to resolve such a uselessly broad question as "does language affect cognition".

A partial list of the genuinely unresolved issues around linguistic relativity are:

(1) What, if any, conceptual or processing domains are susceptible to the influence of linguistic categorization? Presumably, the more fundamental these domains are to cognition and the more their processes are shared with non-linguistic species, the less likely there might be an effect from linguistic categorization. However, is there a discernable boundary between those domains affected by language and those which are not?

(2) What variation exists across languages, i.e., to what extent are languages locked into one way of expression vs. having alternative modes of expression? If there is little variation of a particular category in human language, then it is essentially impossible to decide whether that variation is due to innate cognitive constraints or some other universal guiding principle. It may be that certain language categories are an inevitable feature of a communication system with the complexity and constraints of human language without any need for a genetically predetermined mechanism. Conversely, the features of the studied languages which are said to vary across these languages are often poorly described. For instance, a language may be described as lacking a particular category, when it does in fact express that category but in a way which the researcher did not attend to. The controversy about whether Mandarin expresses conditional reasoning is one such example (cf. the brief discussion of the Bloom controversy below).

(3) Further, what is the nature of linguistic expression needed for there to be a notable effect on non-linguistic cognition? Traditionally, linguistic relativity

studies have focused on grammatically obligatory marking of conceptual categories. When a category is grammatically obligatory, it is taken to be more fundamental than when its expression is optional and less common. On the other hand, when a speaker does express such a category through a more circumlocutory means this certainly implies availability of the category. Pederson et al. (1998) addressed a purported language effect on cognition by constraining both the language and the cognitive task to a specific context. The relationship between language encoding in that context was taken to relate to cognitive processes involved in roughly the same context. What relationship other language uses might have to other cognitive tasks was unexplored.

(4) Similarly, under what conditions do language categories affect non-linguistic cognition (whatever that may be) and when do they not affect cognition? Linguistic relativity studies need not assume that any effect of language categorization on cognition will be uniform and constant. Most narrowly, a study can only argue that an effect is found in the experimental setting used. It is unclear how far one can extrapolate from one experimental setting to a broader class of human behavior. Understandably, opponents of linguistic relativity studies will tend to dismiss any discovered effects as being task-specific while proponents will extrapolate as far as possible.

(5) What is the mechanism of any interaction between language categories and the rest of cognition? For all the advances of cognitive psychology in the past decades, we still lack a detailed model of human cognitive processes. This makes it particularly challenging to formulate specific hypotheses about how one language's categorization might influence other processes. Nonetheless, linguistic relativity studies can still look for behavioral evidence for some effect on non-linguistic processing—even if the exact nature of that process can only be speculative.

26.5 What might a language effect be?

Should research demonstrate a correlation between a variable pattern of language and a corresponding variation in other cognitively driven behavior, there remain a couple of issues with the interpretation of results.

(1) How do we infer from any correlation between a language pattern and a cognitively driven behavior that there is a particular direction of causation? Perhaps the cognitive systems vary across two different cultures for some reason other than language. This most likely would be because of a cultural difference driving the difference in both language use and cognitive patterns across the

two cultures. Such an argument was presented in Li and Gleitman (2002) with a rebuttal in Levinson et al. (2002). That language variation is simply a reflection of non-linguistic cultural variation seems initially a reasonable possibility. Certainly this is the case for recent adoption of lexical items expressing relatively new concepts, i.e., the cultural concern predates the linguistic pattern. On the other hand, if the linguistic pattern in question is represented by a long-fossilized grammatical construction, then it seems far more likely that the linguistic pattern predates any non-linguistic cultural pattern. Of course, since language is one of the major vehicles for transmitting culture across generations, it is not straightforward to distinguish between linguistic and non-linguistic cultural patterns. For further discussion of linguistic relativity vis-à-vis culture, see Hanks (1990), Bickel (2000), and Enfield (2000).

(2) Another concern is that the experimental task seeking to measure a non-linguistic cognitive pattern might not in fact be a non-linguistic task because it is actually mediated by internal processing of language. Should this be the case, a relationship between language and non-linguistic cognition would not be demonstrated. While it is possible to block linguistic behavior during a task by having participants repeat nonsense syllables or engage in other language-masking behavior, this is usually considered too intrusive in a task of appreciable difficulty. The more consciously accessible the desired solution is to the participant, the more likely it is that the participant might choose to adopt a (conscious) strategy of relying on language to solve the task. On the other hand, Pederson (1995) argues that, if a participant selects to use language as the means for solving a task, the participant must understand the categories of that language to be appropriate and reflective of the cognitive categories appropriate to the task. If there were a disconnection between the categories of language and the cognitive categories which would otherwise be used, it is unclear what would sanction the reliance on language.

This last point relates to the often cited "thinking for speaking" notion presented in Slobin (1991) and taken up by many. This idea allows for a language effect on conceptualization for those representations which must be talked about. After all, if the grammar of a language requires encoding a certain type of information, that information should be encoded in the underlying representation when it is to be communicated. One possibility is that all information that any language might require to be encoded would be part of the representation of every speaker of every language. Only the elements required by a particular language would necessarily be brought into focus by speakers of that language, but the universal set of concepts is encoded prior to any linguistic coding. Since the superset of distinctions which any language might require is far greater than what any one language requires, it is clearly more efficient that speakers encode the information which they know will be needed for expression in their particular language and bother less with other information in the absence of other reasons to encode that information. In other words, for the purpose of speaking, speakers will encode events in memory differently depending on the communicative requirements of their languages.

This thinking-for-speaking model appeals as a compromise between heavy-handed linguistic determinism and complete universalism. Unfortunately, it is unclear exactly how such a model might work. If a speaker witnesses an event, will she encode to memory the information needed for retelling only when the event is suspected to be one she will want to communicate about later? This seems unlikely and unworkable. Any event might need to be described later, so conservatively the speaker should always encode the information to be communicated (or habitually invent it on retelling). This then becomes tantamount to speakers encoding information using language as their guide, which is essentially the premise of linguistic relativity and not a compromise position at all. In short, it is unclear how a model of cognition could be built with separate processes of thinking for speaking and thinking for not speaking.

26.6 Language development

Related to linguistic relativity studies with adults is a growing body of research investigating cross-linguistic variation in child development. After all, if linguistic categorization helps to direct cognitive categorization, there must ultimately be a developmental account to explain this. One's first language is notably learned during a period of remarkable conceptual development. In keeping with the universalist bias in psycholinguistics and with a shortage of first language development studies across a diverse set of languages, it was all to easy to assume that children learned language by mapping universal conceptual categories onto a fairly constant set of word meanings. There would certainly be little variation expected in the semantics of child language even if the adult languages seem to exhibit different patterns. The work of Bowerman, Brown, Choi, and others, e.g., Bowerman and Choi (2001), Brown (2001), Choi and Bowerman (1991), and the collection Bowerman and Brown (2008), have challenged these assumptions by arguing that the lexical meaning of children's words are strongly influenced by the idiosyncrasies of the target language. Children's lexical categories can vary dramatically cross-linguistically—suggesting that the process of language-specific category formation can begin early in development. Further, the target languages can differ substantially from one another in precisely the ways that make linguistic relativity questions interesting.

As Whorf is to linguistic relativity, so is Vygotsky (1986) to studies of linguistically mediated child development. In his model, the acquisition of the categories of language is assumed to be a primary vehicle for the development of a (deeper) understanding of these concepts. For example, de Villiers and de Villiers (2000; 2003) credit the acquisition of complementation strategies of natural language as allowing the development of representations of the beliefs of others. Lucy and Gaskins (2001; 2003) have investigated a correlation between Yucatecan and American behavior in a sorting task and the corresponding organization of lexical categories in Yucatec Mayan and English. Interestingly, and broadly consistent with Vygotsky, the acquisition of the language-specific lexical categories appears to precede by a few years the development of the corresponding sorting behavior. This sorting behavior is taken to reflect underlying preferences for conceptual categorization.

26.7 Domains of research

As mentioned, linguistic relativity studies generally explore domains of language and cognition for which one might expect strong universal tendencies. Any such findings will naturally be most vigorously scrutinized and, should they survive such scrutiny, they will be all the more theoretically influential. Accordingly, most linguistic relativity research can be found in just a few cognitively fundamental domains. Even on the most charitable readings for the various linguistic relativity studies to date, some domains have clearly been more fruitful for purported language effects than others, though it is not always clear why this may be so.

Color Color has long been one of the battlegrounds for the linguistic relativity debate. On the one hand, languages clearly vary in their color terminology, suggesting that there may well be substantive differences in speakers' organization of color categorizations. On the other hand, the reference of color terms nonetheless appear subject to some universal/perceptual constraints since, barring color blindness, people have essentially identical color perception prior to any higher-level cognitive processing of that information.

One of the first empirical linguistic relativity studies to be published compared English and Zuni color categorization (Lenneberg and Roberts 1956). Kay and Kempton (1984) is also one of the more classic citations. As an apparent language effect occurred under one condition presentation of color stimuli, but not under a slightly different presentation, Kay and Kempton is frequently cited as support both for and against linguistic relativity effects. Some more recent work on color terms and linguistic relativity includes the substantial collection in Hardin and Maffi (1997). Explicitly expanding the work of Kay and Kempton is the extensive work by

Davies and colleagues (Corbett and Davies 1997; Davies 1998; Davies and Corbett 1997; Davies et al. 1998; Oezgen and Davies 1998) which explores for demonstrable effects of color naming on color categorization tasks.

Space Space has also been assumed to be a domain with a strong universalist underpinning in cognition and in language. Since all humans interact with the same basic environmental properties, there seemed little reason to actually conduct cross-cultural and cross-linguistic investigations in this domain. Seizing on previously little reported, but globally widespread, variation in the linguistic expression of space, Levinson and colleagues have argued that fundamental differences in spatial reasoning and memory encoding co-vary with linguistic expression. Two well-known works in this area are Levinson (1996*a*) and Pederson et al. (1998), which examine languages which habitually use different expressions of reference frames, that is, of the coordinate systems by which one locates objects relative to one another in space. They present spatial memory tasks demonstrating an underlying difference in the categories used to reconstruct spatial arrays from memory.

Unsurprisingly, these arguments have met with criticism from those believing that the human cognition of space must be fairly autonomous from general patterns of language. This skepticism tends to be strongest from those who have not worked with languages differing from the modern European norms. Li and Gleitman (2002) argued that spatial cognition is flexible and largely contextually driven rather than linguistically motivated. They modified the Levinson experiments by manipulating the physical environment of the experiment in a way that they claim demonstrates that the underlying frame of reference is determined by experimental context rather than habitual language encoding. However, the modifications Li and Gleitman made to the experiments are rejected as misleading in Levinson et al. (2002). When the experiments were reconducted to the original specifications with Dutch participants (in Levinson et al. 2002) and with American-English speakers (Church 2005), participants continued to behave in a way better predicted by linguistic pattern than the immediate testing environment. This debate is also continued in Majid (2002) and Majid et al. (2004). Clearly linguistic relativity studies—both arguments for and against a language effect—need to carefully attend to the details of design. For an extensive summary of this work with spatial reference frames consult Levinson (2003*b*).

Time In contrast to investigations into space, there has been little exploration in potential variation in temporal organization. This seems primarily due to three reasons. First, time is essentially one-dimensional, which is clearly simpler than multi-dimensional space, so the logical possibilities of variation are necessarily reduced. Second, at least linguistically, time is generally considered to be expressed in ways derivative of spatial expression (though see Tenbrink 2007 for a critique of this). If the categories of time are essentially derivative of space, then why not study space as the more fundamental domain? Third, and non-trivially, time is generally considered as abstract and metaphorical. While this suggests that variation of expression might be widespread, it also suggests that experiments may be quite challenging to design.

Capitalizing on the metaphorical nature of temporal expression in Mandarin and English, Boroditsky (2000; 2001) argues that different patterns of linguistic expression drive different conceptualizations of time across the two languages. Note, however, that Chen (2007) and January and Kako (2007) fail to replicate her findings and suggest that the linguistic hypothesis itself may not have been adequate. This reinforces the point made above about the importance of careful linguistic description prior to hypothesis.

Motion Combining space and time, we have motion events. Encoding motion events is one of the more fundamental tasks of natural language. Motion events are presumably perceptually universal, yet they are complex enough to suggest a range of conceptualizations should be possible. Talmy (1985) presented a typology of motion events in natural language suggesting that languages type according to one of two basic types in their expression of the path component of a motion event. Slobin and colleagues have used language production data from a variety of languages to argue that the type of language one uses has substantial consequences for how motion events will be represented; see Slobin (2000; 2003) and Slobin et al. (in press). A number of studies have tried to find cognitive correlations in wholly non-linguistic tasks with speakers of different language types with mixed results (Bohnemeyer et al. 2007; Finkbeiner et al. 2002; Naigles and Terrazas 1998; Oh 2003; Papafragou et al. 2001). Loucks and Pederson (in press) argue that this is not so much because there is no language effect to be found in the conception of motion events but that the Talmy typology is insufficient for the purposes of generating a testable hypothesis. Since the processing of motion events is of such fundamental importance and languages do vary in their default representations of such events, we can expect linguistic relativity research to continue in this domain.

Grammatical gender and number Various studies of European languages have observed that speakers (perhaps unsurprisingly) evaluate the references of nouns as having more masculine or feminine qualities based on their (cross-linguistically variable) grammatical gender assignment. For a brief summary of this research, see Boroditsky et al. (2003). This research has not yet expanded to include the influences of nominal categorization across the broader and more varied range found in linguistic typology (see, for example, the collection in Craig 1986). It does seem that this could serve as an interesting domain for future linguistic relativity research.

Noun phrases also vary cross-linguistically as to their expression of plurality. Some languages mark dual as well as plural. Other languages only rarely explicitly mark plural and rely on context to imply a difference in number specification. As cited in the introduction, Lucy (1992*a*) examined grammatical number differences across Yucatecan Mayan and American English speakers and argued that this language difference played out in a difference in the memory encoding of numbers of objects in visual drawing.

Logic and arithmetic number Domains which are susceptible to the influence of formal training have been particularly controversial for linguistic relativity studies. Usually people's abilities to use logic and number are viewed as stronger or weaker rather than as different but equal. This generally requires a hypothesis that one community is advantaged and another disadvantaged by their default linguistic code. Two areas in particular have received particular attention: counterfactual reasoning and arithmetic number.

Bloom (1981) proposed that Mandarin lacks an explicit counterfactual construction and that this leads to a greater challenge in counterfactual reasoning for Mandarin speakers than, for instance, for English speakers who do not lack such a construction. Subsequent studies have argued that while Mandarin may lack a dedicated counterfactual construction, it does have regular means of creating sentences which are clearly counterfactual in context. See the debate spread across Au (1983; 1984), Bloom (1984), Liu (1985), and more recently Cara and Politzer (1993). Further, Lardiere (1992) found that her sample of Arabic speakers patterned more like Bloom's Mandarin speakers than like English speakers despite Arabic having a counterfactual construction. From this, she reasonably concludes that other cultural factors than language are at play.

Miura and colleagues (Miura 1987; Miura et al. 1988; Miura and Okamoto 1989; Miura et al. 1993; Miura et al. 1994; Miura et al. 1999) have argued that speakers of Mandarin, Japanese, and other languages which have a consistent base-ten lexical set are advantaged in learning arithmetic over speakers of languages which have words like "eleven" and similar irregular numbers. It is particularly challenging to factor out family educational values and other cultural factors from such studies. Saxton and Towse (1998) also found that seemingly subtle changes in presentation could make the purported language effect disappear. They interpret this to suggest that any language effect about base-ten numbers is quite indirect.

Watson (1987) argued that the differing grammatical treatment of number in Yoruba disadvantaged monolingual Yoruba children in learning early arithmetic compared to their peers who are bilingual with English. Greiffenhagen and Sharrock (2007) provide a largely philosophical rebuttal against this work as part of a larger argument that linguistic relativity is scarcely an empirical enterprise.

Emotion and personality Impressionistically, there is considerable variation in emotional responses and personality types across cultures. To date, the linguistic descriptions of emotion and personality terms have far too heavy a reliance on translation to allow for testable linguistic relativity hypotheses to be developed. It is nonetheless an area ripe for exploration as it is at least intuitively possible that the categories of emotion and personality expressed in language provide a template along which individuals may mold themselves. See, for example, Marmaridou (2006) for a linguistic description of Greek pain lexicalization.

Working within a single language, Lindquist et al. (2006) find priming and suppression effects in categorization of facial expression from the presentation of words denoting emotions. These findings at least raise the possibility that regular use of language-specific emotion terms may well influence speaker's processing of emotions.

26.8 Summary

All of the research in linguistic relativity to date makes up only the smallest fraction of work within linguistics and the other cognitive sciences. In fact, given the general public interest in the topic, one could say that there has been an appallingly small amount of research. For many, linguistic relativity studies are readily dismissed as counter to current theoretical assumptions. That said, the relationship between the most uniquely human characteristics—linguistic communication and our astounding cognitive capacities—is clearly of profound interest. The last decade has shown a dramatic surge of academic interest in linguistic relativity, and the hypotheses generated and the methods employed to test these hypotheses have shown steady development. After approximately fifty years, the field of linguistic relativity studies may still be young but it shows every sign of developing into an exciting and robust field of research.

CHAPTER 27

RELEVANCE THEORY

FRANCISCO YUS

27.1 Cognition

Relevance theory (henceforth RT) is a cognitive pragmatics theory of human communication which was developed in the mid-1980s by Dan Sperber and Deirdre Wilson (henceforth S&W or, where appropriate, W&S) in their book, *Relevance: Communication and Cognition* (1986, 2nd edition 1995) but their earlier publications (e.g., W&S 1981; see section 27.3 below) also dealt with this theory, specifically comparing it to Grice's cooperative principle. Since then, it has become a highly influential theory in today's pragmatics and has been applied to many types of discourse and research areas within pragmatics (see Yus 1998*a*; 2006).[1]

For S&W, human beings have developed an ability to maximize the relevance of the stimuli[2] that they process. Since it is utterly impossible to pay attention to the entire *barrage* of information that reaches us, we have developed an inherent capacity to filter potentially irrelevant information and to focus our attention on what, in the current situation, is bound to provide cognitive reward. As W&S (2002*a*: 254) state, "as a result of constant selection pressure towards increasing

[1] An extensive online bibliography on relevance theory, arranged in thematic sections and with links to more than 400 downloadable papers can be found in *Relevance Theory Online Bibliographic Service*, at www.ua.es/personal/francisco.yus/rt.html.

[2] A stimulus is any input for mental processing. Stimuli can broadly be divided into verbal stimuli (e.g., an utterance) and nonverbal stimuli (e.g., a gesture).

efficiency, the human cognitive system has developed in such a way that our perceptual mechanisms tend automatically to pick out potentially relevant stimuli, our memory retrieval mechanisms tend automatically to activate potentially relevant assumptions, and our inferential mechanisms tend spontaneously to process them in the most productive way".

Filtering information which does not appear to be relevant (for example, when we do not recall most of the people who pass by us in the street but do remember those who, for some reason, stand out from the crowd) is indeed a typical mental activity oriented toward this relevance-seeking procedure, but it is also essential for human survival to identify underlying intentions and attitudes in the actions (communicative or otherwise) of those who are around us (for example when someone approaches us and we cannot help wondering what intention underlies his or her actions). Besides, human beings tend to select from context only the information that might be useful for obtaining interesting conclusions (contextual information is vast but we have developed a capacity for accessing just the right information that leads to interesting conclusions) and also combine new information with information already stored in their brain or which is accessible at that stage of interpretation (this is essential in human communication for obtaining interesting conclusions; see below). This inherent ability of humans to focus their attention on potentially relevant information is covered in the so-called *cognitive principle of relevance*.

Cognitive principle of relevance:
Human cognition tends to be geared to the maximization of relevance.

This is a biologically rooted principle that is applied to all kinds of processing, including linguistic processing. This is reflected in the general objective of RT: to identify underlying mechanisms, rooted in human psychology, which explain how humans communicate with one another (S&W 1986/95: 32). Among the relevance-oriented tasks undertaken by the human mind, one of the most interesting ones is the human ability to combine contextual or accessible information with new incoming information to yield relevant conclusions, as in (1):

(1) New information (visual input):
A yellow Mercedes is parked near our department.

(2) Information already available (from encyclopedic knowledge):
a. Professor Smith, who supervises my thesis, owns a yellow Mercedes.
b. Professor Smith usually takes the bus to university.
c. Only when he intends to stay at university till late in the evening does he drive his car to university (since there are no late buses returning to where he lives).

(3) (Relevant) conclusion (inferred by combining (1) and (2)):
This evening I will be able to discuss with him at length how my thesis is progressing.

S&W claim that in a situation where (1) is processed, (3) would be relevant since it can only result from the combination of (1) and (2). A similar procedure also applies to linguistic communication (see section 27.2 below), specifically to *intentional* verbal communication. However, relevance is not only applied to external stimuli but also to internal mental representations, some of which are more prominent or likely to be entertained in the current context of interpretation. Consider the following example:

(4) The bell has just rung.

(5) a. Someone has rung the bell.
b. The bell in my house has just rung.
c. The person who is ringing is not a dwarf (he or she can reach the bell).
d. There is no power failure in my building.
e. The company providing electricity has not gone bankrupt.
f. Nobody has stolen my ring.
g. I have paid my latest electricity bill.

In situation (4), some thoughts are more accessible (more *manifest* in RT terminology; see section 27.5 below) and more likely to be entertained than others. In normal circumstances, (5a) and (5b) are the most likely thoughts. However the choice of thoughts is constrained by contextual information. For instance, in a context where there have been a lot of power failures recently, (5d) will then be more relevant and perhaps even more likely to be entertained than other thoughts which would be considered more accessible (manifest) in normal circumstances.

27.2 Communication

The biologically rooted capacity that human beings have developed in order to interact fruitfully with the surrounding world is also applied by S&W to communication, although, in this case, we are dealing with a highly sophisticated tool, language, which helps us transfer thoughts to one another. In a nutshell, whenever someone talks to us, we engage in a relevance-seeking inferential procedure which relies on the so-called *communicative principle of relevance*.

Communicative principle of relevance:
Every act of overt communication conveys a presumption of its own optimal relevance.

When this principle is satisfied (normally, any time anybody addresses us, but also applicable to processing documents such as novels, news items, etc.), addressees undertake an interpretive task which aims at selecting the most appropriate interpretation from the range of interpretations that the utterance (or text) has in the current context. A stimulus has optimal relevance when two conditions are fulfilled. An ostensive stimulus is optimally relevant to an audience only if: (a) it is relevant enough to be worth the audience's processing effort; and (b) it is the most relevant one compatible with a communicator's abilities and preferences (W&S 2002*a*: 256). On paper, hearers will follow the following general procedure:

(a) Follow a path of least effort in constructing an interpretation of the utterance (and in particular in resolving ambiguities and referential indeterminacies, in going beyond linguistic meaning, in supplying contextual assumptions, computing implicatures, etc.).
(b) Stop when your expectations of relevance are satisfied.

And for expectations to be satisfied, the selected interpretation should satisfy two conditions:

(a) An assumption is relevant to an individual to the extent that the positive[3] cognitive effects achieved when it is optimally processed are large.
(b) An assumption is relevant to an individual to the extent that the effort required to achieve these positive cognitive effects is small.

On paper, new information is relevant when it *reinforces* the hearer's assumptions about the world, when it *contradicts* and *eliminates* assumptions and, most importantly, when it *combines* with existing assumptions to generate conclusions (i.e., *implications* or *implicatures*) which cannot be obtained from either this new information or from the existing assumptions taken separately, but only from the combination of both. Consider the following example:

(6) Tom: So ... Did you enjoy going to the cinema last night?
Ann: John was also at the cinema.

For Tom to interpret Ann's utterance correctly as an answer to his question (i.e., about whether she liked going to the cinema or not), he cannot simply interpret Ann's words literally (which apparently have nothing to do with the question) but has to access contextual information (in this case encyclopedic information about Ann), for instance (7a–e) which, when combined with Ann's words, will help Tom reach the intended interpretation (8):

(7) a. Ann has just gotten divorced.
b. Her ex-husband is called John.

[3] Initially, S&W only mentioned "cognitive effects", but in later publications a differentiation was made between those effects which are beneficial (positive cognitive effects) and those which are not. Needless to say, relevance is aimed at positive cognitive effects.

c. Now Ann and her ex cannot stand each other.
d. Whenever they come across each other they argue.
e. Ann gets depressed every time she argues with her ex.

(8) She didn't enjoy going to the cinema last night.

Notice that this example does not differ too much from the one provided in (1)–(3) above. In (1) the new information was visual and it combined with stored encyclopedic information. In (6) the input is linguistic and it is also combined with encyclopedic information to get the right interpretation (i.e., the right conclusion). Crucially, (1) and (6) also differ in the role of intentionality. The former is an interpretation of unintentionally communicated information, whereas (6) involves obtaining an interpretation which Ann *intentionally* wants Tom to process. This is important because pragmatics does not undertake the study of information which reaches the person without a prior intention, although the cognitive mechanism to grasp relevant conclusions applies to both intentional and unintentional communication. Within this picture of intentional (specifically *ostensive*) communication speakers devise their utterances from among certain choices to code their thoughts, and hearers infer which interpretation, from among a choice of possible interpretations in the current context, is the one that the speaker intends to communicate. Wilson (1994: 44) summarizes these basic ideas of RT in four statements: (a) every utterance has a variety of possible interpretations, all compatible with the information that is linguistically encoded; (b) not all these interpretations occur to the hearer simultaneously; some of them take more effort to think up; (c) hearers are equipped with a single, general criterion for evaluating interpretations; and (d) this criterion is powerful enough to exclude all but one single interpretation, so that, having found an interpretation that fits the criterion, the hearer looks no further.

27.3 Grice and the role of intention

Most of the initial research by S&W on relevance was intended to acknowledge the importance of Grice in the history of pragmatics but also to criticize several points of Grice's theory (see Grice 1975), specifically his emphasis on the need for a *cooperative principle* and its *maxims*[4] to explain communication and also his dividing

[4] For Grice, showing a cooperative attitude entailed the fulfillment of several maxims: *maxim of quality* (tell the truth), *maxim of quantity* (provide as much information as necessary), *maxim of relation* (be relevant), and *maxim of manner* (be brief, avoid obscurity of expression, don't be ambiguous). When the speaker's underlying cooperative attitude is taken for granted, the hearer can infer additional information (implicatures) from the fact that he or she is not following any of these maxims, supposedly for a reason.

line between explicitly and implicitly communicated information (see W&S 1981). A summary of S&W's main criticisms of Grice's theory is provided below:

(a) For Grice, understanding an utterance is a matter of constructing the best hypothesis about the speaker's meaning. S&W agree with him on this. But Grice proposes the *cooperative principle*[5] and its maxims as a means of evaluating alternative interpretive hypotheses. For S&W this *principle* is unable to do so.
(b) Similarly, the aim of a pragmatic theory is to explain how hearers identify the speaker's intended interpretation. But Grice does not indicate how the *cooperative principle* would do that, so we cannot explain which interpretations are more likely to be selected by a hearer.
(c) Grice claims that the hearer should take into account the *cooperative principle* when selecting an interpretation but provides no insight into how this is done. Sometimes there are alternative interpretations to be chosen and various implications to be derived. A satisfactory pragmatic theory should be able to explain why hearers choose some interpretations and reject others. Grice's framework does not do this.
(d) For Grice, pragmatic interpretation is an intelligent, inferential process which is based upon conscious reasoning. This is evident in the complex steps he suggested for the derivation of implicatures. But normally people are not really aware that they are inferring interpretations but, rather, engage in an unconscious and spontaneous mind-reading activity. S&W argue that an investigation of relevance helps us to understand why utterances raise the expectations they do, but it also leads them to reject Grice's *cooperative principle* and its maxims and to construct an alternative theory within cognitive pragmatics.

However, S&W also acknowledge Grice's important contribution to pragmatics, especially concerning the importance he gave to the role of intention in communication. Basically, for Grice, understanding an utterance involves recognizing the intentions underlying it. This is an evolved biological predisposition that humans also use in the interpretation of the nonverbal behavior of their interlocutors. What S&W do is to extend this view by proposing two types of intention. On the one hand, the speaker has an *informative intention*, the intention to communicate some information ("a set of assumptions" in the RT terminology). On the other hand, the speaker has a *communicative intention*, the intention to alert the interlocutor of his or her informative intention. Crucially, successful communication demands

[5] "Our talk exchanges do not normally consist of a succession of disconnected remarks, and would not be rational if they did. They are characteristically, to some degree at least, cooperative efforts; and each participant recognizes in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction" (Grice 1975: 45). The definition of his *cooperative principle* is as follows: "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged."

the fulfillment of both intentions (in this case it is called *ostensive* communication). This is particularly evident in the case of nonverbal stimuli. Consider the following situation:

(9) Tom and Ann are at a disco. At a certain stage during the night, Ann feels very tired and wants to go home. She sees Tom at the other end of the disco. He also sees her and waves at her. Since there is no way he can hear her at that distance, she moves her arms as if she was driving a car so that Tom can infer that she wants him to give her a ride home.

In situation (9) it is of the utmost importance for Tom to infer that Ann is intentionally producing her nonverbal behavior in order to inform him of something, otherwise Tom might think that Ann is simply engaged in some personal form of dancing. Only by inferring Ann's intentionality will Tom understand her correctly. In RT terms, only by inferring Ann's *communicative intention* can Tom understand the right *informative intention* ("I want you to take me home").

The valuable quality of verbal communication lies in how effectively it satisfies the speaker's communicative intention. Indeed, when a person speaks to us we immediately infer that the person is willing to communicate some information to us, and therefore the communicative intention is immediately satisfied and we can focus our inferential activity on the message—informative intention—that the speaker intends to communicate. Verbal communication is indeed an invaluable tool for transferring thoughts.

27.4 Coding/inferring

Traditional linguistic theories had a rather simplistic view of how people communicate information to one another. In short, for these theories all that speakers have to do is to *code* their thoughts into words and send them through a channel (e.g., the air, a book). Hearers are supposed to perform the same tasks but in reverse order: receive the words from the channel and *decode* them into the speaker's thoughts. Duplication of thoughts between speakers and hearers was, from this point of view, the norm, rather than the exception. This is the *code model of communication*.

Within RT, by contrast, the picture is much more realistic and the emphasis is laid upon inference rather than on coding. Although thoughts have to be coded in order to be transmitted to other people, the information that speakers literally code is far more limited than the information that they really intend to communicate with this coded message. In other words, the utterance normally *underdetermines* the information that is intended and this gap between what is coded and what is intended is filled by inference (see Carston 2002).[6] This is why the RT model of interpretation is called the *inferential model*. This model predicts that natural language sentences do not encode propositions but schemas for the construction of propositional forms. This idea is based on the evidence that we cannot possibly code *literally* the thoughts that we entertain, so instead of the uselessly lengthy (10b) a person should in normal circumstances utter a more relevant (10a):

(10) a. I think Susan put the book on one shelf in the sitting-room.
b. I think Susan Thomas, my sister, put the book by Sperber that she bought two days ago on one of the shelves that are located in the sitting-room of this house, downstairs.

Under this *underdeterminacy thesis* (the claim that what people literally say is less informative than what they really want to communicate), there are two types of informational resemblance in human communication with gaps which have to be filled inferentially:

The interpretation that the speaker intends to communicate with his or her utterance
[is more informative than ...]
the literal meaning of what the speaker says, and this literal meaning
[is less informative than ...]
the interpretation chosen by the hearer.

Typical examples of "informational filling" are provided in italics in the following examples frequently found in the bibliography on this issue:

(11) I haven't eaten.
I haven't eaten [*this morning*].

(12) It will take time to fix your car.
It will take [*longer than you'd expect*] to fix your car.

(13) Everybody left early.
Everybody [*at the party*] left early.

(14) There's nothing on TV tonight.
There's nothing [*worth watching*] on TV tonight.

[6] In this aspect, S&W also differ from Grice. For Grice, there is a very small gap between what is said and what is explicitly communicated, with reference assignment and disambiguation as the only pragmatic tasks undertaken to obtain the explicit interpretation. At the same time, he put all the inferential load into the derivation of implicatures, his famous coining. The work by RT analysts such as Carston (2002) has shown that obtaining explicit interpretations can be as demanding (or even more so), in terms of contextualization, as obtaining implicatures.

27.5 Mutuality

Within pragmatics it is commonly assumed that for communication to be successful, there has to be some information which is shared by both interlocutors, and on which they rely when communicating their thoughts. For example, for analysts of *pragmatic presupposition*, there is always some information supposedly shared by both interlocutors when an utterance is articulated and communicated successfully, as in (15):

(15) a. A: Where is Tom?
B: There is a yellow BMW outside Sue's house.

b. Pragmatic presupposition (supposedly shared): Tom owns a yellow BMW.

c. Implication (warranted by (b)): Tom must be at Sue's.

Although it is intuitively certain that people do share information when they talk to each other, S&W reject this traditional notion of *mutual knowledge* because it generates an endless recursion (A knows that p, B knows that A knows that p, A knows that B knows that A knows that p, and so on) that prevents us from really assessing what is truly shared. Instead, they propose the notion of *mutual manifestness* (see S&W 1990). What is "manifest" is what one is capable of inferring or perceiving at a certain stage of interpretation, even if one has not yet done so. The sum of all the manifest assumptions is the person's *cognitive environment*, which varies from one context to another.[7] A set of assumptions manifest to several individuals constitutes their *shared cognitive environment*. This, again, changes according to different contextual parameters. When it is manifest to all the people sharing a cognitive environment that they share it, this is a *mutual cognitive environment*, made up of mutually manifest assumptions. Communication, then, is a matter of making certain assumptions mutually manifest to both speaker and hearer.

27.6 The semantics/pragmatics distinction

There is currently a lot of scholarly discussion on where to place the dividing line between semantics and pragmatics. Within RT the semantics/pragmatics

[7] Which means that manifestness is a matter of degree: in a certain context some assumptions will be more manifest than others, and hence more likely to be processed. In example (5) above, the thought that someone has rung the bell is *more manifest* than the thought that the person ringing is not a dwarf, even if both thoughts are manifest.

distinction is related to the coded/inferred distinction. When we interpret a person's utterance, it is first apprehended by the language module of the brain which delivers a schematic *semantic representation* or *logical form* of the utterance. The "language module" terminology comes from the *modularity thesis*, since S&W initially follow Fodor's (1983) modular picture of the brain, according to which the mind is made up of a (mysterious) central processor, capable of an immense number of computations, and of a number of modules which "feed" the central processor with information. Modules are evolved, special-purpose mental mechanisms, typically automatic and informationally encapsulated. One of these modules is the language module, which is only (and automatically) activated by verbal stimuli, feeding the central processor with a schematic logical form. As summarized in Yus (2006: 516), over the last few years, this view of the mind has changed within RT especially concerning the structure of the central processor, which is also considered to be modular. The most important module in this central processor, specifically a sub-module of the general "theory of mind" ability, is *the pragmatic module*, which also exhibits qualities typically associated with modules. For example, this pragmatic module is biologically endowed, only activated by a specific type of information (ostensively communicated information), and constrained by its own principle: the *communicative principle of relevance*.

The context-free logical form is then enriched inferentially in order to obtain a fully contextualized interpretation (see section 27.8 below). In this picture, semantics would be in charge of the context-free semantic representation of the utterance, whereas the inferred interpretation would belong to pragmatics. In Blakemore's (2002: 23) words,

> The point of contact between semantics and pragmatics is at the interface between the linguistic parser, which receives input from linguistic competence and delivers linguistically determined semantic representations, on the one hand, and the inferential mechanisms which take these semantic representations as input for the computations which deliver the representations which the hearer takes to be representations of the speaker's communicative intentions, on the other.

27.7 Relevance as a cost–benefit procedure

Relevance is measured by hearers by following a cognitive cost–benefit procedure. All things being equal, the hearer will tend to select the interpretation, from the range of possible interpretations of the same utterance in the current context,[8] that satisfies the conditions of highest reward (positive cognitive effects) and least mental effort, although hearers will normally be willing to devote extra cognitive effort if they are going to get additional effects.

As mentioned above, information which *reinforces* previous assumptions, *contradicts* and *eliminates* previous assumptions, or *combines* with previous or accessible assumptions to obtain conclusions provides the highest number of cognitive effects. Given a range of choices, the hearer is entitled to select the one providing the highest interest. But interest is constrained by effort. Wilson (2002*a*) provides a clear example of how effort constrains the choice of an interpretation:

> Imagine exactly the same information being presented, first in a clearly printed form; second as a faint photocopy; third as an illegible handwritten scrawl; fourth translated into a language you read only with difficulty. Each of these versions may have exactly the same cognitive effects for you, but each will require different amounts of processing effort. Although they carry exactly the same information, you will have to work harder to retrieve it from one input than from another, and this may affect your intuitions of relevance, and indeed, your willingness to attend to a particular input at all. More generally, an input may be more or less perceptually salient, more or less legible, more or less linguistically or logically complex, and may therefore cause more or less effort.

Notice, though, that this cost–benefit procedure does not imply that hearers invariably choose the most effort-relieving interpretation. In practice, they will gladly devote cognitive resources to more effort-demanding interpretations if the eventual reward, in terms of positive cognitive effects, is worth the effort. The same applies to speakers when devising their utterances. An example of how speakers do not always opt for the most economical utterance is provided below:

(16) Ann: (1) Does Susan eat meat?
Tom: a. She is a vegan.
b. No. She doesn't eat meat.

In theory, reply (a) does not provide a direct answer to (1), and produces a higher processing effort for (a) than does a more straightforward answer like (b). The explanation for the choice of a more costly answer such as (a) is that it provides additional interest (cognitive effects) that could not be obtained from (b) (in this case the reason for her refusal to eat meat), and this interest makes up for the increased effort. Consider now the following ad (cf. Tanaka 1994):

(17) Less bread. No jam (ad by London Transport).

According to RT, hearers follow two basic steps during interpretation: (a) they consider interpretations in the order of accessibility; and (b) stop when their

[8] Indeed, information can be relevant in one context and not in another, so the basic notion they want to define is that of *relevance in a context*. By a "context" they mean information ("a set of assumptions" in RT terminology) used in interpreting (or "processing") a new piece of information, either verbal or nonverbal.

expectation of relevance is satisfied. In the case of the ad in (17), the first accessible interpretation is "London Transport is offering something that involves less bread and no jam, probably some type of food". This interpretation is not relevant in this context (actually, the reader is bound to be puzzled by the combination of "food" and "transportation"). The reader will then consider a second possible interpretation of the ad: "less bread" colloquially means "less money"; hence "no jam" refers to "no traffic jams". London Transport offers a service which costs less and involves no traffic jams. This second interpretation *is* relevant in the context of the processing of the ad. Moving to the second interpretation has meant more mental effort for the reader, who is nevertheless satisfied with finding the interpretation. This satisfaction compensates for the effort.

27.8 The explicit/implicit distinction (in utterance interpretation)

The general two-step procedure for the interpretation of utterances mentioned above is, in reality, a complex cognitive procedure involving a *mutual parallel adjustment* of three sources of information: (a) the explicit interpretation of the speaker's utterance (it has to be enriched in order to obtain a fully contextualized proposition); (b) the speaker's implicated interpretation—implicature—(if intended); and (c) the right amount of contextual information needed to get (a) and (b). Unlike Grice's *two-step* model of communication according to which one first interprets the utterance literally, concludes that this interpretation is not possible, and then moves on to the implicit or implicated interpretation, S&W predict a dynamic and flexible human cognition capable of accessing context, enriching the utterance at the explicit level, and deriving implicated conclusions without a fixed order, and only constrained by our inherent search for relevance. The parallel subtasks for interpretation are summarized in (18):

(18) a. Construct appropriate hypotheses about explicit content (*explicatures*) via disambiguation, reference assignment, and other pragmatic enrichment processes (see below).
b. Construct appropriate hypotheses about the intended contextual assumptions (*implicated premises*).
c. Construct appropriate hypotheses about the intended contextual implications (*implicated conclusions*).

As pointed out above, interpretation starts with the identification, by the language module, of the schematic and context-free logical form of the utterance which has to be enriched inferentially. This logical form is turned into a fully contextualized proposition called *explicature*.[9] To turn the logical form into an explicature (subtask (18a)) some inferential operations have to be performed (depending on the inferential requirements of the utterance):

— *Reference assignment* and *free enrichment*. Sometimes a referent has to be found for certain words in the utterance. This is typically the case of utterances containing indexicals (i.e., pronouns, adverbs, etc.). Free enrichment, on the other hand, is the inferential completion of the propositional content of the utterance which, despite being apparently complete, needs extra information (e.g., unarticulated constituents) to make sense, as in the bracketed additions in these examples (provided by Carston 2001):

(19)	a. Paracetamol is better.	[than what?]
	b. It's the same.	[as what?]
	c. He is too young.	[for what?]
	d. It's hot enough.	[for what?]
	e. I like Sally's shoes.	[shoes in what relation to Sally?]

This inferential "completion" process is obligatory if the hearer wants to make sense of the intended interpretation and without it there would be no relevant propositional form or explicature.

— *Disambiguation*. When the utterance contains a polysemous word, one of its senses has to be selected according to contextual constraints.

— *Conceptual adjustment*. This is one of the most interesting lines of research within RT. During interpretation, the concept coded by a word is adjusted by the hearers so that it meets their expectations of relevance. The outcome is an *ad hoc concept*[10] which is similar but not identical to the stabilized concept coded by the word.

In certain contexts, the concept that the speaker intends to communicate is *broader* (less exact) than the concept that the word he or she has chosen literally communicates, as in (20a–e):

(20) a. There is *a rectangle* of lawn in the shed.
[not an exact rectangle]

[9] Relevance theory's explicature/implicature distinction is as follows: An assumption communicated by an utterance U is explicit if and only if it is a development of a logical form encoded by U. On the analogy of *implicature*, S&W (1986: 182) call an explicitly communicated assumption an *explicature*. Any assumption communicated, but not explicitly so, is implicitly communicated: it is an *implicature*. Besides explicatures, there are also *higher-level explicatures*, which include the speaker's attitude (*to regret that*..., *to be happy that*..., etc.) or a higher-order speech-act schema (*to be asking that*..., *to be ordering that*... etc.).

[10] They are ad hoc "because they are not linguistically given, but are constructed online in response to specific expectations of relevance raised in specific contexts. There is a difference then between ad hoc concepts, accessed by a spontaneous process of pragmatic inference, and lexicalized concepts, which are context-invariant" (Carston 2002: 322).

b. We entered a pub, but we left since it was *empty*.
[there were people in the pub—e.g., the waiter—but not interesting people]
c. I've got *a thousand things* to do this morning.
[many things, but not a thousand]
d. Don't worry. I'll be ready *in two minutes*.
[in a while, surely longer than two minutes]
e. This steak is *raw*.
[not literally raw, but undercooked]

On other occasions, the concept that the speaker intends to communicate is *narrower* (more exact) than the concept that the word he or she has chosen literally communicates,[11] as in (21a–e):

(21) a. I've got *nothing* to wear for the party.
[nothing appropriate, nothing classy, etc.]
b. María has *a brain*.
[not simply a brain: an outstanding brain; she is very intelligent]
c. This boy has *a temperature*.
[a higher temperature than he should have]
d. It will take *some time* to fix this car.
[longer than you imagine; longer than it would normally take]
e. Antonio *drinks* too much.
[drinks too much alcohol]

The notion of *ad hoc concept* is particularly interesting for the analysis of metaphors. The relevance-theoretic account of metaphors is based on the assumption that there is an *interpretive resemblance* between concepts. There is a difference, though, between the initial RT approach and the current one. Indeed, although both accounts rely on the notion of "interpretive resemblance", in the initial approach the relation was between the concept (or thought) of the speaker and the propositional form of the utterance, and in the new account it is between an encoded concept and a concept communicated.

Besides, both accounts involve the derivation of a range of (strong and/or weak) implicatures.[12] The only major difference is that in the second an ad hoc concept occurs in the explicature, thereby giving inferential warrant to the implicatures derived. In other words, it is claimed that the metaphor provides a new *ad hoc* concept for the proposition expressed by the utterance (which *is* communicated

[11] According to Carston and Powell (2005: 283), "while most other pragmatic approaches assume that narrowing and broadening are to be treated as distinct processes, the RT view is that they are simply different possible outcomes of a single pragmatic process which fine-tunes the interpretation of virtually every word".

[12] Implicatures can be stronger or weaker depending on the amount of contextual assumptions that the hearer needs to retrieve in order to obtain them.

as an explicature). This concept is part of the information that the hearer uses to derive metaphoric implicatures.

Typically, metaphors involve *both* broadening and narrowing of concepts, as in utterance (22a) with the intended interpretation (22b), which demands both broadening (22c) and narrowing (22d) (example from Wilson and Carston 2006):

(22) a. My daughter is *a princess*.
b. My daughter, who is not a female royal, is a spoiled, overindulged girl, who constantly asks for special treatment, expects her wishes to be granted, refuses to do housework, etc.
c. The ad hoc concept is *broader* than the encoded concept in some respects since it applies to some people who are not actual princesses.
d. The ad hoc concept is also *narrower* in some respects since it applies only to people—including princesses—who are spoiled, overindulged, etc.

In the aforementioned *mutual parallel adjustment* of explicit content, implicit import, and access to contextual information, the inferential tasks of reference assignment, disambiguation, etc. are applied, when necessary, to the logical form in order to develop the proposition which is communicated explicitly (*explicature*) but, at the same time, this proposition is combined with the right amount of context to yield, again if necessary, fully inferential implicated premises and implicated conclusions (*implicatures*), and all of these inferential tasks are guided by our biologically rooted search for relevance. Let's exemplify this with the exchange in (23):

(23) Tom: So ... Did you buy that table I told you about?
Ann: It's too wide and uneven.

If Tom wants to understand Ann correctly he has to use inference in order to develop the schematic logical form provided by Ann's utterance into a relevant interpretation. Some inference will be devoted to obtaining the contextualized propositional form of the utterance which is communicated as an explicature. In this particular case, Tom has to engage in reference assignment ("it" refers to "the table"), disambiguation (a table can be "uneven" in several ways: because its surface is uneven or because its legs are not properly leveled), and free enrichment (e.g., too wide [for what?]). The outcome could perhaps be the proposition in (24):

(24) Explicature: "The table that you told me about is too wide to go through the bedroom door and its surface is uneven."

This is not the actual answer to Tom's question, so Tom also has to combine (24) with contextual information (*implicated premises*) in order to get the intended interpretation (*implicated conclusion*). In this case *encyclopedic* contextual information will be accessed by Tom about how unlikely it is for a person to buy a table that does not go through the door and whose surface is uneven. This contextual information will help Tom reach, as an implicature, the intended interpretation (25):

(25) Implicature: "I didn't buy the table that you told me about" (*implicated conclusion*).

27.9 Relevance and grammar

Most of the studies of grammar which take RT as the theoretical framework move beyond the traditional view of grammar to a more dynamic and inference-centered approach in which grammatical senses are not taken for granted but supported or refuted according to contextual constraints. In short, grammatical aspects are no longer intrinsic and stable features of language, nor are grammatical attributes a mere list of choices in hypothetical contexts supplied by the grammarian. Instead, a pragmatic and consequently context-centered view of grammar is proposed in which grammatical attributes constrain (or not) the choice of a right (i.e., intended) interpretation. The addressee's ability to access the adequate context in which the utterance can be optimally processed also plays an important part in the outcome of interpretation. In this case, the grammatical organization of utterances has an important role throughout this cognitive contextualization, since it often imposes constraints upon the range of possible interpretations of the utterance and thus reduces (or increases) the effort required to select the intended interpretation (Yus 1997: 237). A short review of studies which, one way or another, analyze grammatical aspects of language using RT is provided below.

(a) *The conceptual/procedural distinction*. This is one of the most important contributions of RT to the study of grammar, which has resulted from the work by Blakemore (e.g., 1987; 2002) on connectives and discourse markers and has led to a great number of studies in the same area. Instead of a typical approach to connectives, which tends to make a basic distinction between a same-level relationship of elements (coordination, parataxis) and a hierarchy-based one (subordination, hypotaxis), connectives including *after all*, *so*, *but*, *whereas*, etc. are regarded as constraints on relevance, that is, as guidelines for the correct comprehension of the compound sentence, since they reduce the effort needed to access the correct interpretation: "[T]heir sole function is to guide the interpretation process by specifying certain properties of context and contextual effects. In a relevance-based framework, where the aim is to minimise processing costs, the use of such expressions is to be expected" (Blakemore 1987: 77). This minimization of effort can take different directions, since these terms aid in obtaining the speaker's intended effects by restricting the construction of either the explicatures or the implicatures of the utterance.

Blakemore (2002: 82) lists three possible attributes that we can expect an expression which encodes procedural meaning to have. Firstly, there is "elusiveness", in the sense that procedural expressions are hard to paraphrase or translate and their descriptions are usually controversial (for instance, there is a difference between the procedural *but* and the non-procedural *in contrast*, the latter being easier to paraphrase).

Secondly, procedural discourse markers do not have synonymous VP adverbial counterparts. One of the examples by Blakemore (2002: 84) shows how *in other words* is used in the same sense in (26) and (27), whereas *well* is used differently in (28) and (29), which implies that *in other words* encodes conceptual meaning and *well* encodes procedural meaning:

(26) *In other words*, you're banned.

(27) She asked me to try and put it *in other words*.

(28) A: What time should we leave?
B: *Well*, the train leaves at 11.23.

(29) You haven't ironed this very *well*.

Thirdly, elements that encode conceptual meaning can be semantically complex whereas elements that encode procedural meaning cannot.

Consequently, from this perspective, connectives such as *so*, *but*, and *after all* are used in order to make implicit coherence relations explicit, and hence to establish a safe guideline for the interpretation of utterances containing them. For instance, the connective *but* helps the hearer to infer that the proposition it introduces is relevant as a denial of the expectation or as a contrast regarding the proposition expressed in the first clause.

W&S (1993: 10) argue that "inferential comprehension involves the construction and manipulation of conceptual representations; linguistic decoding feeds inferential comprehension; linguistic constructions might therefore be expected to encode two basic types of information: concepts or conceptual representations on the one hand, and procedures for manipulating them on the other". Connectives such as *but* should not be seen as encoding concepts but as procedural devices which *constrain* the inferential phase by indicating the kind of cognitive process that the hearer should go through (hence reducing the eventual overall effort).

In further research on the conceptual/procedural distinction the list of expressions encoding procedural meaning has been extended to other elements of language, for example punctuation marks (Borochovsky Bar-Aba 2003), tense, mood (see below), and even to nonverbal qualities of communication such as intonation (Escandell-Vidal 1998).

(b) *Conditionals*. In general, there is some discussion on the semantic vs. pragmatic uses of conditionals. Smith and Smith (1988) and other authors claim that the RT framework is useful to clarify this point. They suggest that the behavior of both factual and counterfactual conditionals can be explained in contextual, relevance-theoretic terms. For example:

(30) a. If you are confident enough, bet your whole salary on that horse.
b. If I ask you politely, will you post the letter?
c. If you are hungry, there is a flan in the fridge.

In these sentences the hearer has to recover the propositional form of the sentence (via enrichment of the logical form) and integrate it into a description according to the imperative (*the speaker is telling the speaker to p*, as in (30a)), or an interrogative connotation, as in (30b), and in both cases there is a guarantee of relevance for the speaker and/or hearer. For (30c), Smith and Smith propose the following RT-related explanation:

> [...] the antecedent specifies a state of affairs which, as usual, provides a relevant context for the consequent. Given the Principle of Relevance, this in turn forces the listener to make certain additional assumptions: specifically, that he can infer from the guaranteed relevance of the consequent that the flan in the fridge is available for him. Given the easily accessible information that hunger is undesirable, that eating alleviates hunger and that flans are for eating, the force of the whole conditional is accounted for naturally. (Smith and Smith 1988: 335)

(c) *Modals and modality*. Several authors have addressed modals and modality using a relevance-theoretic approach (e.g., Berbeira Gardón 1996, Groefsema 1995, and Nicolle 1997). They attempt to provide a cognitive explanation for the various senses in which modals can be used in similar contexts. For example, *may* has both *epistemic* and *deontic* interpretations of (31a) in (31b) and (31c) respectively:

(31) a. She may do the examination tomorrow.
b. It is possible that she will do the examination tomorrow.
c. She is permitted to do the examination tomorrow.

Under RT, modals are considered to have a basic meaning, and the different interpretations which they can acquire are dependent on contextual attributes. In other words, there is a basic propositional meaning which is later enriched to yield a propositional form with a context-related (epistemic or deontic) meaning. Groefsema (1995: 61) goes on to say that "the basic meanings of *can, may, must* and *should* express relations between the proposition expressed by the rest of an utterance containing them and a set of 'background' assumptions, while putting constraints on what sets of assumptions are recovered during the interpretation process".

(d) *Adverbs and adverbials*. Ifantidou-Trouki (1993), among others, has studied adverbs and adverbials under RT. She deals with four types of adverbials: illocutionary (*frankly, confidentially, honestly* ...), attitudinal (*unfortunately, happily, sadly* ...), evidential (*evidently, obviously* ...), and hearsay (*allegedly, reportedly*). They are usually regarded as indicators of the type of speech act performed with the utterance. The RT-based analysis proves that in reality these kinds of adverbials are very different from each other in their use. Indeed, not all adverbs have a procedural role in the utterances where they occur, helping the hearer in their processing, but they can also encode conceptual representations.

(e) *Mood*. Clark (1993), W&S (1988), and Ahern and Leonetti (2004), among others, have studied moods such as subjunctive and imperative. Clark's concern is with pseudo-imperatives. These include verbs used with a covert conditional meaning, as in (32a–d) and also the kind of imperative sense of *let's* in (32e):

(32) a. Wash the car and I'll buy you an ice-cream.
b. Leave the house or I'll call the police.
c. Come one step closer and I'll shoot.
d. Turn on the radio and you'll hear the news about the murder.
e. Let's go to the movies tonight.

Clark claims that RT can explain their grammatical behavior, basically proposing that the semantic content of the utterance is combined with contextual information in order to access its intended interpretation. In this sense, the explanation of the "conditional" interpretation of sentences such as (32a–d) is entirely pragmatic: "[I]n each case the hearer has to make some assumption about how desirable the state of affairs is thought to be and from whose point of view it is thought to be desirable; in making these assumptions he is guided by contextual factors and considerations of optimal relevance" (Clark 1993: 82).

On the other hand, Ahern and Leonetti (2004: 37) argue that in the case of verbal mood, its semantic content contributes to the specification of explicatures, mainly those known as higher-level explicatures (*attitude*- or *speech-act*-connoted propositions), in which the speaker's propositional attitude and communicative intention are represented. This conceptualization of mood fits the aforementioned procedural/conceptual dichotomy, since "the semantic content of the grammatical moods is minimal compared to the range of interpretive effects that the use of one or the other can convey: their stable, unitary semantic content leads to a variety of diverse interpretive effects depending on the context they are used in" (ibid.).

(f) *The article*. Jucker (1992) unifies within the RT framework the different senses of the definite article that are typically proposed in grammars. Hearers can only make hypotheses about their interlocutors' assumptions, and successful communication depends on these assumptions being accurate. This, when applied to noun phrases, implies that speakers continuously wonder if their interlocutors will manage to identify the intended referent of the noun phrase or not, with the following sub-assumption: "The referent of the expression which contains a definite article is uniquely identifiable at the particular point of the discourse at which the expression occurs" (ibid., 128), which is part of the meaning of the definite article and therefore underlies most of the uses proposed by the grammar: "[T]he various categorisations of uses of the definite article that have been proposed by grammarians in essence try to compartmentalise the bases on which speakers assume that their addressees will be able to pick up the correct referent" (ibid., 130).

(g) *Tense*. Authors such as Smith (1990) and Haegeman (1989) suggest that RT provides an ideal framework to avoid the ambiguities that arise in logical descriptions of tense. Indeed, for RT it is not enough to say that the past tense provides a temporal reference prior to the moment of speaking. Instead, the hearer is expected to "narrow the reference down to some more specific interval, so that the utterance can be constructed as expressing an optimally relevant proposition; that is, so that it can interact with accessible contextual assumptions to give rise to a range of effects" (Smith 1990: 85). Smith goes on to apply RT to the sequentiality of the narrative past in sentences such as the following:

(33) John entered the office. The president walked over to him.

in which the information provided by the first part precedes that provided by the second part. Smith concludes that this is part of "reference assignment" and consequently part of the propositional enrichment to yield an explicature of the sentence.

Haegeman's analysis focuses on the difference between *going to* and *will* in an attempt to overcome the intuitive distinction of these auxiliaries in examples such as (34a–b):

(34) a. I will/shall leave next week.
 b. I'm going to leave next week.

Her claim is that at the level of sentence meaning, the meaning of both auxiliaries is equivalent, and that the difference is to be found in the constraints that they impose on their processing in the context of the utterance in which they occur. In her study she concludes that,

> *Be going to* . . . imposes a constraint on the processing of the proposition with which it is associated. It signals that this proposition is relevant in a context including at least some present tense propositions, or, in other words, it guarantees a contextual effect if the utterance is processed against a present context. *Will*, on the other hand, signals that the hearer should extend the immediately accessible (present) context for the processing of the proposition and should process the utterance against future propositions. (Haegeman 1989: 305)

(h) *Aspect*. Žegarac (1990; 1993) analyzes aspect from a relevance-theoretic approach. Starting off with a traditional classification of aspect and situation types, he concludes that under RT these traditional labels seem problematic. For example, in the case of stative verbs, which cannot be used in the progressive aspect, many instances are found which are clear exceptions to this rule (1990: 127):

(35) a. Peter is being polite.
b. Mary is loving the fruit salad.

In this case, the hearer realizes that (35a) and (35b) are not intended to be interpreted literally, and starts looking for assumptions about what the speaker might have intended to convey (normally in the form of implicatures). Basically, the heart of the matter lies in the quality of the encyclopedic entries for the verbs concerned and the hearer's accessibility to contexts in which the stative meaning of the verbs is eventually created. Some examples by Žegarac (1990: 129) are particularly illustrative:

(36) a. John doesn't feel well.
b. John isn't feeling well.

(37) a. The baby resembles her mother.
b. The baby is resembling her mother more and more.
c. ?? The baby is resembling her mother.

(38) a. Antoinette understands Russian.
b. Antoinette is understanding Russian better and better.
c. ?? Antoinette is understanding Russian.

For Žegarac, the meaning of the progressive is to be defined in terms of reference to instantiation(s) of the property denoted by the predicate, and the predicates in (36–38) take the progressive. However, "feel well" does so more readily than the predicates "resemble one's mother" and "understand Russian", acceptable only in (37b) and (38b), which contain explicit indications of change ("more and more", "better and better"). The difference between the progressive form of "feel", on the one hand, and "resemble" and "understand Russian" on the other, would be accounted for by RT not as a difference in the degree of "stativity" or "dynamicness" inherent in the meanings of these verbs but as a difference in the accessibility of contexts which are used to achieve adequate contextual effects.

27.10 Social aspects of communication

RT has been criticized for avoiding the social aspects of communication. Recent discussion includes Rajagopalan's (2005) claim that S&W cannot account for intercultural aspects of communication. S&W (1997: 147) acknowledge that they have concentrated on the inferential activity of the individual, but inferential communication is also essentially social: "Inferential communication is intrinsically social, not just because it is a form of interaction, but also, less trivially, because it exploits and enlarges the scope of basic forms of social cognition. Right or wrong, this is a strong sociological claim." In W&S (2005: 100) they add: "[A]lthough pragmatists generally see communication as both a cognitive and a social process, they do not always devote their efforts equally to developing rich accounts of both the cognitive and the social factors involved. We see this as a difference in interests and research strategies rather than in theoretical commitments. In our own work, we have focused on cognitive factors, but we still assume that a comprehensive picture of communication should integrate both kinds of factors."

As pointed out in Yus (2006), a proposal by Escandell-Vidal (2004) aims at integrating individual and social issues in terms of principles and norms, respectively, and as part of a dynamic picture of human inference. The mind operates according to principles that are in charge of obtaining fully propositional interpretations from coded stimuli. When dealing with norms, the mind is engaged in both a long-term and a short-term task. The latter analyzes and categorizes incoming information, and the former builds up and updates socially accepted behavior.

27.11 Empirical evidence

A common criticism of RT is that it is highly speculative, predicting without empirical evidence the mental procedures and interpretive steps the human mind goes through in human communication. Obviously, we are dealing with an object of study, the human mind, that is highly complex and still largely unexplained. S&W (2002: 143) acknowledge that in much pragmatic research there is a certain reluctance to get down to experimentation. But recent research on a number of pragmatic issues has shed light onto empirical evidence for RT. For instance, Van der Henst and Sperber (2004) review various experimental tests of the two *principles of relevance*. Other studies have focused on *Wason selection task*. The plausibility of a Gricean maxim of truthfulness to explain human communication has also been tested. The research showed that when people ask a stranger the time in the street they get, as a reply, "a time that is either accurate to the minute or rounded to the nearest multiple of five, depending on how useful in the circumstances they think a more accurate answer would be" (W&S 2002*b*: 598), regardless of whether the people asked have (accurate) digital watches. These rounded answers are not strictly true, but more relevant in that context.

27.12 Relevance and other areas of pragmatic research

RT has been applied to many research areas such as *humor* (e.g., Yus 2003), *media discourses* (e.g., Tanaka 1994; Yus 1998*b*, 2001), *literature* (e.g., Pilkington 2000), *politeness* (e.g., Jary 1998), *irony* (e.g., W&S 1992), *translation* (e.g., Gutt 2000), and *language teaching* (e.g., Nizegorodcew 2007) in addition to grammar. These research areas which take RT as their theoretical framework are clear evidence of the dynamism and impact of this cognitive pragmatics theory of communication.

CHAPTER 28

ROLE AND REFERENCE GRAMMAR AS A FRAMEWORK FOR LINGUISTIC ANALYSIS

ROBERT D. VAN VALIN, JR.

28.1 Introduction: Motivation, goals, and evidence

There are many motivations for proposing and developing a theoretical framework in linguistics. For example, a leading idea in the development of Generalized Phrase Structure Grammar (Gazdar et al. 1985) was to determine whether natural language syntax could be adequately described in terms of a context-free phrase structure grammar. Lexical-Functional Grammar [LFG] (Bresnan 1982*a*, 2001) had a number of motivations, including applying the formalism of unification grammar to natural language phenomena and showing that lexical rules were superior to transformational rules. Construction Grammar (Fillmore 1988; Goldberg 2006) emerged as a reaction to Chomsky's claim that constructions are mere epiphenomena derived from the interaction of more basic general rules and principles. Role and Reference Grammar [RRG] was inspired by both typological and theoretical concerns. The motivating questions for RRG were, "what would a linguistic theory look like if it were based on the analysis of languages with diverse structures, such as Lakhota, Tagalog, Dyirbal, and Barai (Papua New Guinea), rather than on the analysis of English?", and "how can the interaction of syntax, semantics, and pragmatics in different grammatical systems best be captured and explained?" These two questions contain both theoretical and descriptive content. On the one hand, they both emphasize the importance of taking account of typologically diverse languages in the formulation of a linguistic theory, and, on the other, they indicate that the resulting theory will be one in which semantics and pragmatics play significant roles. One of the reasons that some of the constructs posited by RRG are rather different from those in other theories is precisely because of this starting point. For example, theories starting from English and other familiar Indo-European languages often take the notion of subject for granted, whereas for one that starts from syntactically ergative and Philippine languages, this is not the case, and the notion of subject as a theoretical construct is called seriously into question. Indeed, the initial work in RRG concerned the question of the universality of subject and the cross-linguistic validity of grammatical relations in general (Foley and Van Valin 1977, 1984; Van Valin 1977*a*, 1981).

Since the late 1980s there have been two additional questions that RRG seeks to answer: "can language acquisition be accounted for without recourse to an autonomous Language Acquisition Device?", and "can a model of grammar that answers the typological and theoretical questions posed above provide any insights into the neurocognitive processing of language?" The final chapter of Van Valin and LaPolla (1997) (henceforth VVLP) is devoted to the first question, and the tentative conclusion is that the acquisition of quite a few core grammatical phenomena can be explained in RRG terms, including some for which it has been argued that there is no evidence available to the child regarding them, e.g., subjacency.[1] In the last few years there has been increasing work on applying RRG to language processing, both in computational and neurolinguistic terms. Computational implementation of RRG is in its infancy (Kailuweit et al., 2003, Butler 2004, Nolan 2004, Guest 2008, Guest and Brown 2007), but the results so far are very promising; in particular, Guest (2008) has developed a parser based on RRG and has used it to successfully parse a large corpus of English sentences as well as a small corpus of Dyirbal sentences, including ones with discontinuous constituents. With respect to sentence processing, one of the distinctive features of RRG, to be discussed in more detail below, is the bidirectionality of the mapping between syntax and semantics.

[1] In addition to the references in Van Valin and LaPolla (1997) see also Van Valin (1998, 2001*b*, 2002), Weist (2002), Weist et al. (2004).

The RRG linking algorithm maps from semantics to syntax and from syntax to semantics. This is an idealization of what a speaker does (semantics to syntax) and what a hearer does (syntax to semantics). Hence the design of the theory makes it readily amenable to psycho- and neurolinguistic sentence processing models. Recently there has been experimental neurolinguistic research involving RRG as the grammar component of a sentence comprehension model (Bornkessel, Schlesewesky and Van Valin 2004, Bornkessel and Schlesewsky 2006*a*, 2006*b*, Van Valin 2006, Bornkessel-Schlesewsky and Schlesewsky 2008). The success and explanatory power of this model can be taken as support for the RRG linking system.

An important issue with respect to the questions RRG seeks to answer is the nature of the evidence that can be brought to bear on them, and in this regard the theory is particularly eclectic. Obviously, data from grammatical descriptions from a wide range of languages is central to the evaluation of analyses and theoretical claims. However, other types of evidence are used as well. The analysis of a number of grammatical phenomena has been influenced by evidence from language acquisition and neurolinguistic experiments. Sociolinguistic and conversational data have also played a role in RRG analyses; the data investigated in Belloro (2007), for example, include the results of a sociolinguistic survey of use of clitic pronouns in Buenos Aires Spanish, and Shimojo (1995) presents an account of *wa* and *ga* in Japanese based on the analysis of transcripts of Japanese TV talk shows. His 2005 book explores argument encoding in Japanese conversation. Thus, RRG takes a variety of evidence into account.

The vast majority of work in RRG has been synchronically oriented, although as far back as Foley and Van Valin (1984) and Ohori (1992) it was argued that there were diachronic implications of RRG analyses. In the last few years there has been more diachronic work, e.g., Matasović (2002), Wiemer (2004), Eschenberg (2004, 2005), and the most recent results are presented in a collection of papers on applying RRG to issues in language change and grammaticalization (Kailuweit et al. 2008).

The presentation of RRG will proceed as follows. In the next section, the organization of theory will be discussed, and the basics of the representations posited in the theory will be introduced. In the subsequent section, the linking system will be described and applied to some much discussed grammatical phenomena. The final section contains a brief summary.

28.2 Organization and representations

The organization of RRG is given in Figure 28.1; it will be elaborated further in § 28.3.

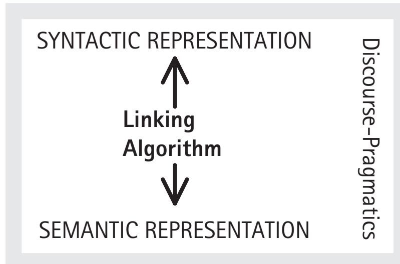


Figure 28.1. The organization of RRG

There is a direct mapping between the semantic and syntactic representations of a sentence, unmediated by any kind of abstract syntactic representations; this excludes not only derivational representations, as in, for example, the Minimalist Program, but also the use of abstract f-structures as in LFG. There is only a single syntactic representation for a sentence, and it corresponds to the actual form of the sentence. RRG does not allow any phonologically null elements in the syntax; if there's nothing there, there's nothing there.[2] RRG posits a very concrete syntactic representation, and this constrains the theory significantly; as noted above, this rules out derivations and therewith movement rules, however they are formulated, and also abstract representations like relational networks in Relational Grammar or f-structures in LFG. Many of the descriptive and theoretical devices in theories with abstract syntactic representations are not available in RRG. The syntactic representation will be described in §28.2.1.

The syntactic representation is linked via the linking algorithm (§28.3) to the semantic representation. It consists of a lexical decomposition representation of the meaning of the predicator along with its arguments. This will be discussed further in §28.2.2.

The last component in Figure 28.1 is labeled "discourse-pragmatics", and it is parallel to the linking algorithm. What this depicts is the fact that discourse-pragmatics, primarily as realized in information structure, plays a role in the linking between syntax and semantics. Crucially, however, exactly what role it plays can vary across languages, and this variation is the source of important cross-linguistic differences among languages. Information structure and its representation will be introduced in §28.2.3.

There is no RRG-related theory of phonology. Work on an RRG theory of morphology is in its initial stages (Everett 2002, Martin Arista 2008), and the representation of the internal structure of words would be part of the syntactic representation. The role of the lexicon will be discussed in §28.2.2 and §28.3.

[2] RRG does allow zero morphemes in morphological paradigms; what is excluded are phonologically null pronouns (*pro*, PRO), noun phrases (traces), light verbs, adpositions, etc., in syntactic representations.

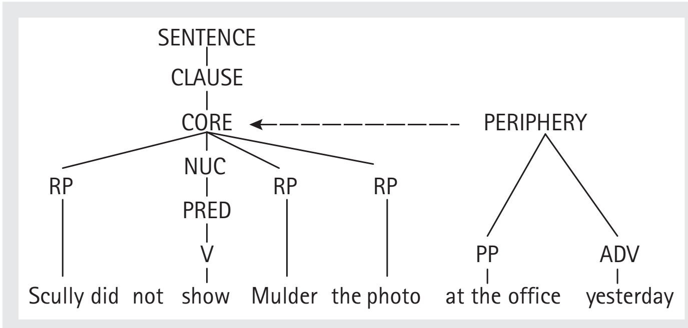


Figure 28.2. The layered structure of the clause in English

28.2.1 The syntactic representation of a sentence

The RRG theory of syntactic representation strives to satisfy the two conditions in (1).

(1) General considerations for a theory of clause structure:
   a. A theory of clause structure should capture all of the universal features of clauses without imposing features on languages in which there is no evidence for them.
   b. A theory should represent comparable structures in different languages in comparable ways.

Clause structure is not represented in RRG in terms of X-bar syntax or even traditional immediate constituency structure;[3] rather, it is captured in a semantically-based model known as the "layered structure of the clause". The essential components of this model of the clause are (i) the NUCLEUS, which contains the predicate, (ii) the CORE, which contains the nucleus plus the arguments of the predicate in the nucleus, and (iii) a PERIPHERY for each layer, which contains adjunct modifiers. These aspects of the layered structure are universal. The structure of a simple English clause is given in Figure 28.2; the structure in Figure 28.2 is the constituent projection of the clause.[4] In Table 28.1 the semantic units underlying the layered structure of the clause are summarized.

The distinctions in Table 28.1 are universal and follow from the fact that language is used to refer and predicate. There is no verb phrase in the layered structure because it is not universal. Hence the layered structure of the clause meets the condition in (1a).

[3] See VVLP, chapter 2, for arguments against a traditional or X-bar phrase structure analysis.

[4] Abbreviations: ABS "absolutive", ADV "adverb", AUH "Actor-Undergoer Hierarchy", DET "determiner", ERG "ergative", IF "illocutionary force", LDP "left-detached position", LOC "locative", LS "logical structure", NMR "non-macrorole", NUC "nucleus", PERF "perfect", PrCS "pre-core slot", PRED "predicate", PRO "pronoun", PROG "progressive", PSA "privileged syntactic argument", RP "reference phrase", TNS "tense".

Table 28.1. Semantic units underlying the syntactic units of the layered structure of the clause

Semantic Element(s)	Syntactic Unit
Predicate	Nucleus
Argument in semantic representation of predicate	Core argument
Non-arguments	Periphery
Predicate + Arguments	Core
Predicate + Arguments + Non-arguments	Clause (= Core + Periphery)

The predicate in the nucleus need not be a head, nor is it restricted to a particular category. While the most common category for the predicate in the nucleus is undoubtedly verb, adjectives, as in *Pat is* ***tall*** (*be* is analyzed as a tense-carrying auxiliary, not as a predicate), nominal phrases, as in *Mary is* ***a very good lawyer***, and adpositional phrases, as in *Sam is* ***at the office***, can all serve as the predicate in the nucleus. Hence the RRG theory of constituent structure is non-endocentric.[5]

Some languages have a "pre-core slot", which is the position of WH-words in languages like English and Icelandic, and a "left-detached position", which is the position of the pre-clausal element in a left-dislocation construction. In addition, some verb-final languages have a "post-core slot" (e.g., Japanese; Shimojo 1995), and some languages also have a "right-detached position", which is the position of the post-clausal element in a right-dislocation construction.

A second important component of the RRG theory of clause structure is the theory of OPERATORS. Operators are closed-class grammatical categories like aspect, negation, tense, and illocutionary force. An important property of operators is that they modify specific layers of the clause. This is summarized in Table 28.2.

Languages normally do not have all of these operators as grammatical categories; the absolutely universal ones are illocutionary force and negation. Operators are represented in a separate projection of the clause, which is the mirror image of the constituent projection. This is exemplified in Figure 28.3.

An example of an English sentence with constituent and operator projections is given in Figure 28.4.

The sentence in Figure 28.4 involves a left-detached position as well as a pre-core slot housing a WH-expression. Note that there is no empty argument position in the core corresponding to the WH-word in the PrCS; see fn. 2. In this example, *did* is labeled both "tense" and "IF" in the operator projection, because the position of the tense operator signals illocutionary force in English: core-medial tense signals declarative IF, pre-core tense signals interrogative IF, and the absence of tense in a matrix core signals imperative IF.

[5] See Everett (2008) for a discussion of intentional state constructions in Wari', an Amazonian language, in which whole clauses can serve as the nucleus of a sentence.

Table 28.2. Operators

Nuclear operators:
- Aspect
- Negation
- Directionals (only those modifying orientation of action or event without reference to participants)

Core operators:
- Directionals (only those expressing the orientation or motion of one participant with reference to another participant or to the speaker)
- Event quantification
- Modality (root modals, e.g., ability, permission, obligation)
- Internal (narrow scope) negation

Clausal operators:
- Status (epistemic modals, external negation)
- Tense
- Evidentials
- Illocutionary Force

An important claim made in Foley and Van Valin (1984) was that the linear order of the morphemes expressing the operators is a function of their scope.[6] That is,

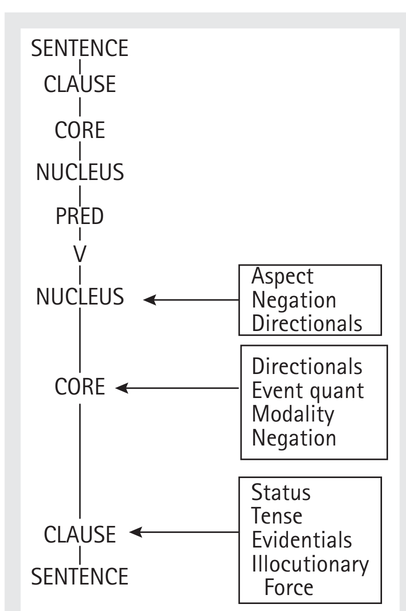


Figure 28.3. The operator projection in the layered structure of the clause

[6] This claim was supported by the results of the typological survey reported in Bybee (1985*a*).

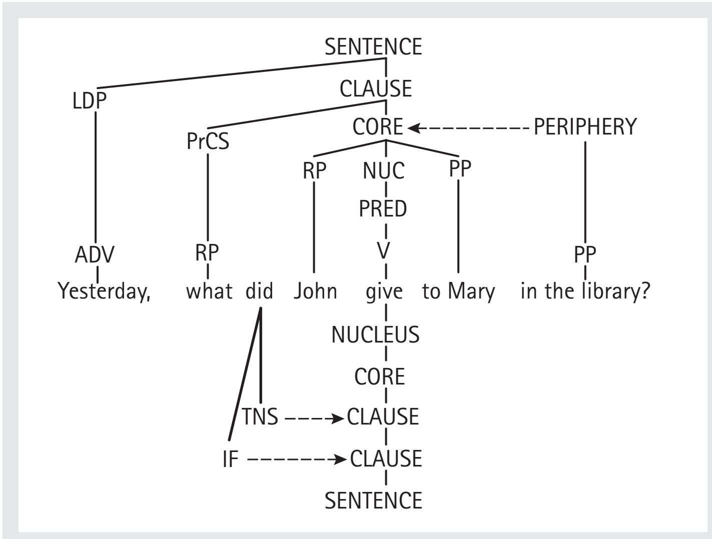


Figure 28.4. An English sentence with both constituent and operator projections

morphemes expressing nuclear operators appear closer to the nucleus than morphemes expressing core operators, and these in turn occur closer to the nucleus than morphemes expressing clausal operators, when an ordering relationship among them can be established, i.e., they all occur on the same side of the nucleus. This is illustrated for two left-branching and two right-branching languages in Figure 28.5.

Other types of phrases also have a layered structure analogous to the clause, and they may have both constituent and operator projections, as appropriate. In Figures 28.2 and 28.4 the nominal phrases are labeled "RP" instead of "NP". "RP" stands for "reference phrase", and unlike "NP" but like the clause, it is a non-endocentric construct. The nucleus of an RP is neither restricted to nominals, nor is it restricted to lexical heads. The first point is particularly important, given the much discussed issues raised by languages like Nootka (Swadesh 1939, Jakobsen 1979) and Tagalog (Schachter 1985, Himmelmann, 2008), in which the heads of referring expressions need not be nominal in nature. See Van Valin (2008*a*) for detailed discussion.

The idea of a layered structure is applied to other categories as well, especially RPs and PPs. RPs have both constituent and operator projections, with the operator projection containing categories such as definiteness, deixis, quantification, and number. Examples of RPs and their layered structure are give in Figure 28.6. This approach to the structure of RPs makes possible an analysis of discontinuous constituency that satisfies principle (1b). The Dyirbal and English sentences in Figure 28.7 are translations of each other, and the Dyirbal example involves discontinuous RPs. The lines connecting the determiners to the head nouns are the operator projection within the RP, as in Figure 28.6. In head-marking languages like Lakhota, the bound pronominals on the verb are considered to be the core

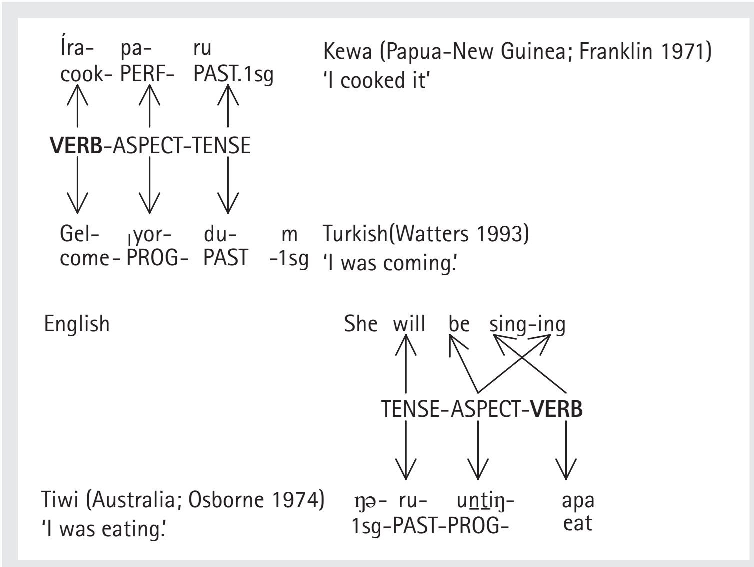


Figure 28.5. Examples of the ordering of aspect and tense markers in different languages

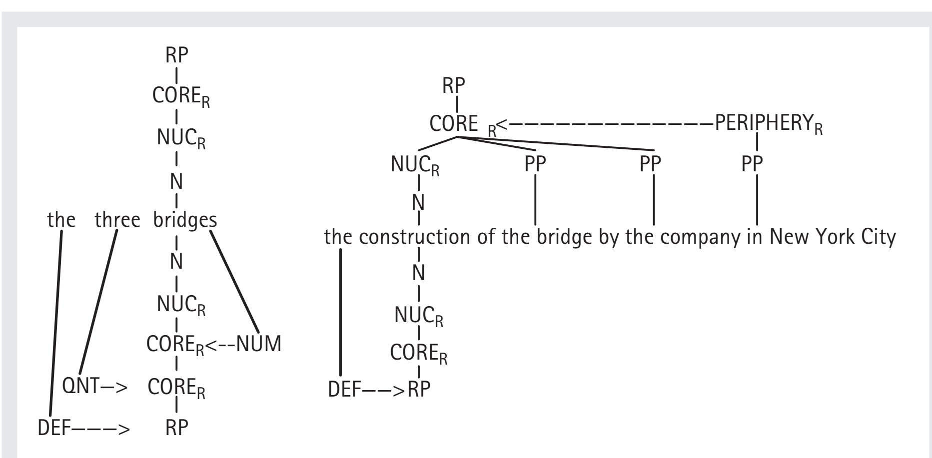


Figure 28.6. The layered structure of the RP with operator projection

arguments; overt RPs are within the clause in apposition to them (Van Valin 1977*b*, 1985). Hence, in the first diagram in Figure 28.8, which means "I killed the bears", the bound pronominals plus the nucleus, all one phonological word, constitute the core of the clause, while the independent RP is clause-internal (it can be the focus of a question, and therefore is within the scope of the IF operator, namely the clause). In the second one, which means "I killed them", the verb word constitutes the core, clause, and sentence by itself. It contrasts with its English translation only in terms

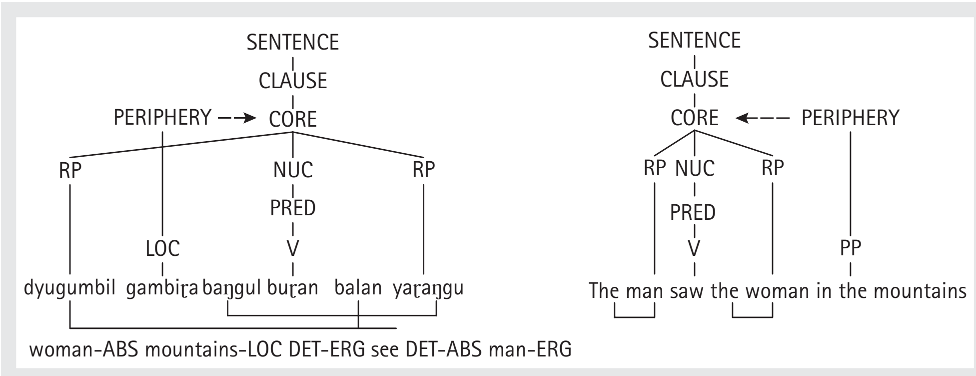


Figure 28.7. The layered structure of the clause in Dyirbal and English

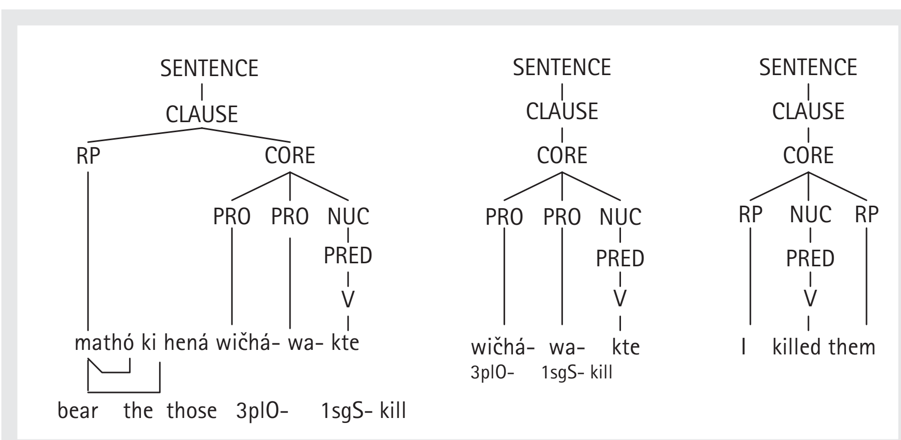


Figure 28.8. The LSC in Lakhota (head-marking) and English (dependent-marking)

of the order of morphemes and the fact that the English pronominals are free rather than bound morphemes.

Note that despite the differences between the three languages in Figures 28.7 and 28.8, comparable structural relations, e.g., core argument, peripheral adjunct, are represented in the same way.

Representations of constituent projections such as these are analyzed as "syntactic templates", the inventory of which in a language constitutes an important component of its grammar. It is termed the "syntactic inventory" and complements the lexicon. There are template selection principles, based on the semantic representation of a sentence, that determine the selection of the proper template(s). The syntactic structure for a sentence may be made up of a combination of multiple templates.

28.2.2 The semantic representation of a sentence

The semantic representation of a sentence is based on the lexical representation of the verb or other predicating element. It is a decompositional representation based on Vendler's (1967) theory of *Aktionsart*. The four basic classes (state, achievement, accomplishment, and activity) are augmented by two additional classes, semelfactives (punctual events; Smith 1997) and active accomplishments (telic uses of activity verbs, e.g., *devour, run to the store*); in addition, there are causative versions of each. Examples of the six classes are given in (2), and sentences illustrating the classes plus their causative counterparts are given in (3).

(2) a. States: *be sick, be tall, be dead, love, know, believe, have*
  b. Activities: *march, swim, walk* (– goal PP); *think, eat* (+ mass noun/bare plural RP)
  c. Semelfactives: *flash, tap, burst* (the intransitive versions), *glimpse*
  d. Achievements: *pop, explode, shatter* (all intransitive)
  e. Accomplishments: *melt, freeze, dry* (the intransitive versions), *learn*
  f. Active accomplishments: *walk* (+ goal PP), *eat* (+ quantified RP), *devour*

(3)	a.	State:	The boy fears the dog.
	a′.	Causative state:	The dog frightens/scares the boy.
	b.	Achievement:	The balloon popped.
	b′.	Causative achievement:	The cat popped the balloon.
	c.	Semelfactive	The light flashed.
	c′.	Causative semelfactive	The conductor flashed the light.
	d.	Accomplishment:	The ice melted.
	d′.	Causative accomplishment:	The hot water melted the ice.
	e.	Activity:	The dog walked in the park.
	e′.	Causative activity:	The girl walked the dog in the park.
	f.	Active accomplishment	The dog walked to the park.
	f′.	Causative active accomplishment:	The girl walked the dog to the park.

Syntactic and semantic tests determine the *Aktionsart* of a clause (see VVLP §3.2.1; Van Valin (2005) (henceforth VV05); §2.1.1). As the sentences in (3e–f′) show, a single verb, e.g., *walk*, can have more than one *Aktionsart* interpretation. This verb would be listed in the lexicon as an activity verb, and lexical rules would derive the other uses from the basic activity use (see VVLP §4.6; Van Valin, forthcoming).

The system of lexical decomposition builds on the one proposed in Dowty (1979). Unlike Dowty's scheme, the RRG system treats both state and activity predicates as basic. The lexical representation of a verb or other predicate is termed its LOGICAL STRUCTURE [LS]. State predicates are represented simply as **predicate**′, while all activity predicates contain **do**′. Accomplishments, which are durative, are distinguished from achievements, which are punctual. Accomplishment LSs contain

Table 28.3. Lexical representations for *Aktionsart* classes

Verb Class	Logical Structure
STATE	**predicate′** (x) or (x, y)
ACTIVITY	**do′** (x, [**predicate′** (x) or (x, y)])
ACHIEVEMENT	INGR **predicate′** (x) or (x, y), or
	INGR **do′** (x, [**predicate′** (x) or (x, y)])
SEMELFACTIVE	SEML **predicate′** (x) or (x, y), or
	SEML **do′** (x, [**predicate′** (x) or (x, y)])
ACCOMPLISHMENT	BECOME **predicate′** (x) or (x, y), or
	BECOME **do′** (x, [**predicate′** (x) or (x, y)])
ACTIVE ACCOMPLISHMENT	**do′** (x, [**predicate′**$_1$ (x, (y))]) & BECOME **predicate′**$_2$ (z, x) or (y)
CAUSATIVE	α CAUSE β, where α, β are LSs of any type

BECOME, while achievement LSs contain INGR, which is short for "ingressive". Semelfactives contain SEML. In addition, causation is treated as an independent parameter which crosscuts the six *Aktionsart* classes, hence the twelve classes in (3). It is represented by CAUSE in LSs. The lexical representations for each type of verb in (3) are given in Table 28.3.

Examples of simple English sentences with the LS of the predicate are presented in (4).

(4) a. STATES
- Leon is a fool. — **be′** (Leon, [**fool′**])
- The window is shattered. — **shattered′** (window)
- Fred is at the house. — **be-at′** (house, Fred)
- John saw the picture. — **see′** (John, picture)

b. ACTIVITIES
- The children cried. — **do′** (children, [**cry′** (children)])
- The wheel squeaks. — **do′** (wheel, [**squeak′** (wheel)])
- Carl ate snails. — **do′** (Carl, [**eat′** (Carl, snails)])

c. SEMELFACTIVES
- The light flashed. — SEML **do′** (light, [**flash′** (light)])
- John glimpsed Mary. — SEML **see′** (John, Mary)

d. ACHIEVEMENTS
- The window shattered. — INGR **shattered′** (window)
- The balloon popped. — INGR **popped′** (balloon)
- John glimpsed the picture. — INGR **see′** (John, picture)

e. ACCOMPLISHMENTS
- The snow melted. — BECOME **melted′** (snow)
- The sky reddened. — BECOME **red′** (sky)
- Mary learned French. — BECOME **know′** (Mary, French)

f. ACTIVE ACCOMPLISHMENTS	
Carl ate the snail.	**do**′ (Carl, [**eat**′ (Carl, snail)]) & BECOME **eaten**′ (snail)
Paul ran to the store.	**do**′ (Paul, [**run**′ (Paul)]) & BECOME **be-at**′ (store, Paul)
g. CAUSATIVES	
The dog scared the boy.	[**do**′ (dog, Ø)] CAUSE [**feel**′ (boy, [**afraid**′])]
Max broke the window.	[**do**′ (Max, Ø)] CAUSE [BECOME **broken**′ (window)]
The cat popped the balloon.	[**do**′ (cat, Ø)] CAUSE [INGR **popped**′ (balloon)]
Bill flashed the light.	[**do**′ (Bill, Ø)] CAUSE [SEML **do**′ (light, [**flash**′ (light)])]
Felix bounced the ball.	[**do**′ (Felix, Ø)] CAUSE [**do**′ (ball, [**bounce**′ (ball)])]
The girl walked the dog to the park.	[**do**′ (girl, Ø)] CAUSE [**do**′ (dog, [**walk**′ (dog)]) & BECOME **be-at**′ (park, dog)]

Full semantic representations of sentences also contain lexical representations of the RPs, adjuncts, and grammatical operators like tense and aspect; see VVLP §4.4, 4.7; VV05 §2.2–2.3.

28.2.2.1 *Semantic macroroles and lexical entries for verbs*

The semantic interpretation of an argument is a function of its position in the LS of the predicate, and, as will be seen below, the linking system refers to an element's LS position. Thematic relations as such play no role in the theory; the traditional thematic role labels are used only as mnemonics for the LS argument positions, e.g., "theme" is the mnemonic for the second position (y) in a two-place locational LS like **be-at**′ (x, y). RRG posits two generalized semantic roles or SEMANTIC MACROROLES, which play a crucial role in the linking system. The two macroroles are ACTOR and UNDERGOER, and they are the two primary arguments of a transitive predication; the single argument of an intransitive predicate can be either an actor or an undergoer, depending upon the semantic properties of the predicate. The basic distinction is illustrated in the following German examples.

(5) a. Der Junge [SUBJ, ACTOR] hat den Kuchen [OBJ, UNDERGOER] aufgegessen.
'The boy ate the cake.'
b. Der Hund [SUBJ, ACTOR] ist um das Haus herumgelaufen.
'The dog [SUBJ, ACTOR] ran around the house.'

c. Der Hund [SUBJ, UNDERGOER] ist gestorben.
'The dog [SUBJ, UNDERGOER] died.'

d. Der Kuchen [SUBJ, UNDERGOER] wurde vom Jungen [ACTOR] aufgegessen.
'The cake [SUBJ, UNDERGOER] was eaten by the boy [ACTOR].'

In (5a), *der Junge* "the boy" is the actor and *den Kuchen* "the cake" is the undergoer of the transitive verb *aufessen* "eat up"; in the sentences with intransitive verbs, *Der Hund* is an actor with the activity verb *herumlaufen* "run around" and an undergoer with the accomplishment verb *sterben* "die". Actor is not equivalent to syntactic subject, nor is undergoer equivalent to syntactic direct object, as the examples in (5c) and crucially (5d) show: in both of these sentences the syntactic subject is an undergoer, and in the passive sentence in (5d) the actor is an oblique adjunct. In an English clause with an active voice transitive verb, the actor is the initial RP (the traditional subject) and the undergoer, when it occurs, is always the direct RP immediately following the verb. In an English passive construction, the undergoer is the subject and the actor, if it occurs, is in an adjunct PP in the periphery$_{\text{CORE}}$.

Actor and undergoer are generalizations across specific semantic argument types, as defined by LS positions. This is illustrated in Figure 28.9.

The *x* argument of all of these verbs functions as the actor, regardless of whether it is the first argument of the generalized activity verb **do**′ (conventionally labeled "effector"), as with *kill, put* and *present*, or the first argument of a two-place state predicate, as with *see*. With two-place transitive verbs like *kill* and *see*, the *y* argument is the undergoer. With three-place verbs like *put* and *present* (as in *Bill presented Mary with the flowers*), on the other hand, the situation is potentially more complex, and this will be discussed in §§ 28.3 and 28.4.

The relationship between LS argument positions and macroroles is captured in the Actor–Undergoer Hierarchy [AUH] in Figure 28.10. The basic idea of the AUH is that in a LS the leftmost argument in terms of the hierarchy will be the actor and the

kill	[**do**′ (x, Ø)] CAUSE [BECOME **dead**′(y)]
see	see′ (x, y)
put	[**do**′ (x, Ø)] CAUSE [INGR **be-LOC**′ (y, z)]
present	[**do**′ (x, Ø)] CAUSE [INGR **have**′ (y, z)]

Actor Undergoer

Figure 28.9. Macroroles as generalizations over specific thematic relations

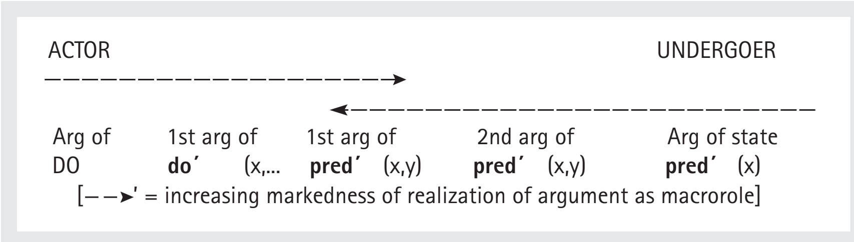


Figure 28.10. The Actor–Undergoer Hierarchy[7]

rightmost will be the undergoer. This was true for *kill, see*, and *put* in Figure 28.9. It was not true for *present*, however, and this reflects a fundamental asymmetry in the AUH: the leftmost argument in a LS (in terms of the AUH) is always the actor, but the rightmost argument is only the default choice for undergoer. This possible variation in the selection of the undergoer is the basis of the RRG analysis of dative shift and related phenomena (see §28.3).

Transitivity in RRG is defined semantically in terms of the number of macroroles a predicate takes. This is termed "M-transitivity" in RRG, following Narasimhan (1998), in order to distinguish it from the number of syntactic arguments a predicate takes, its "S-transitivity". The three M-transitivity possibilities are: transitive (2 macroroles), intransitive (1 macrorole), and atransitive (0 macroroles). It is important to point out in the context of this discussion of three-place predicates that there is no third macrorole; there is nothing in RRG corresponding to Primus' (1999) notion of "proto-recipient". From theoretical and empirical perspectives, there are no grounds for positing a third macrorole; see Van Valin (2004), VV05: 64–6, for detailed discussion). The theoretical label for the third argument in a ditransitive predication; e.g., *the picture*, in the English sentence *Sam showed Sally the picture*, is "non-macrorole direct core argument".

The principles determining the M-transitivity of verbs are given in (6).

(6) Default Macrorole Assignment Principles

a. Number: the number of macroroles a verb takes is less than or equal to the number of arguments in its LS.
   1. If a verb has two or more arguments in its LS, it will take two macroroles.
   2. If a verb has one argument in its LS, it will take one macrorole.

[7] RRG treats the notion of "agent" rather differently from other theories. The basic notion is "effector", which is the first argument of **do**′ and is unspecified for agentivity. With many verbs, a human effector may be interpreted as an agent in certain contexts. If the verb lexicalizes agentivity, as with *murder*, then the logical structure contains "DO", which indicates that the argument must be interpreted as an agent. See Holisky (1987), Van Valin and Wilkins (1996), VVLP §3.2.3.2, for detailed discussion. Also, primary-object languages patterns require a modified undergoer selection principle, namely that the undergoer is the second-highest ranking argument in the LS; see Guerrero and Van Valin (2004), Van Valin (2005: 123–7).

b. Nature: for predicates which have one macrorole,
   1. If the verb LS contains an activity predicate, the macrorole is actor.
   2. If the predicate has no activity predicate in its LS, it is undergoer.

If a verb is irregular and has exceptional transitivity, it will be indicated in its lexical entry by "[MRα]", where "α" is a variable for the number of macroroles. Examples of lexical entries for some English verbs are given in (7).

(7)	a.	*kill*	[**do**′ (x, Ø)] CAUSE [BECOME **dead**′ (y)]
	b.	*receive*	BECOME **have**′ (x, y)
	c.	*own*	**have**′ (x, y)
	d.	*belong (to)*	**have**′ (x, y) [MR1]
	e.	*see*	**see**′ (x, y)
	f.	*watch*	**do**′ (x, [**see**′ (x, y)])
	g.	*show*	[**do**′ (w, Ø)] CAUSE [BECOME **see**′ (x, y)]
	h.	*run*	**do**′ (x, [**run**′ (x)])
	i.	*drink*	**do**′ (x, [**drink**′ (x, y)])

A major claim in RRG is that no syntactic subcategorization information of any kind is required in the lexical entries for verbs. For regular verbs, all that is required is the LS and nothing more, as in all except (7d). For most irregular verbs, only the macrorole number needs to be specified. The prepositions that mark oblique arguments with verbs like *show* are predictable from general principles and need not be listed in the lexical entry (see below, also Jolly 1993; VVLP §7.3.2). All of the major morphosyntactic properties of verbs and other predicates follow from their LS together with the linking system.

28.2.3 The information structure representation of a sentence

The morphosyntactic means for expressing the discourse-pragmatic status of elements in a sentence is called "focus structure", and the approach to focus structure used in RRG is based on Lambrecht (1994). He proposes that there are recurring patterns of the organization of information across languages, which he calls "focus types". The three main types are presented in (8), with data from English and Italian; focal stress is indicated by all caps.

(8) Focus structure in English and Italian (Lambrecht 1994)

a. Q:	What happened to your car?	**Predicate Focus**
A:	i. My car/It broke DOWN.	English
	ii. (La mia macchina) si è ROTTA.	Italian
b. Q:	What happened?	**Sentence Focus**
A:	i. My CAR broke down.	English
	ii. Mi si è rotta la MACCHINA.	Italian

c. Q:	I heard your motorcycle broke down.	**Narrow Focus**
A:	i. My CAR broke down.	English
	ii. Si è rotta la mia MACCHINA./ È la mia MACCHINA che si è rotta.	Italian (Lit: 'broke down my car'/'it's my car which broke down')

Predicate focus corresponds to the traditional topic-comment distinction, with a topical subject RP and a focal predicate phrase which receives the focal stress. It is universally the least marked or default focus structure. In English, the subject would most likely be an unstressed pronoun, while in Italian it would most likely not occur at all; if it were overt, it would be preverbal in Italian. Sentence focus is a topicless construction in which the entire sentence is focal. In English, the subject receives the focal stress, while in Italian the subject appears postverbally and with focal stress. Narrow focus involves focus on a single constituent, in these examples, the subject. In English this is signaled by focal stress on the element or by a cleft, e.g., *It was my* CAR *that broke down*. Italian likewise has two options: postposing the subject, when it is the focused element, or a cleft.

There is an important distinction between unmarked and marked narrow focus. All languages have an unmarked focus position in the clause; in English it is the last constituent of the core, whereas in verb-final languages it is the position immediately before the verb. Consider the following English sentence with different focal stress options.

(9) a. Dana sent the package to LESLIE yesterday.
b. Dana sent the package to Leslie YESTERDAY.
c. Dana sent THE PACKAGE to Leslie yesterday.
d. Dana SENT the package to Leslie yesterday.
e. DANA sent the package to Leslie yesterday.

Focal stress on *Leslie* in (a) is a case of unmarked narrow focus, while focal stress on any other constituent of the clause, as in (b)–(e), yields marked narrow focus. The most marked narrow focus is on the subject, as in (e).

Information structure is represented by an additional projection of the clause, the focus structure projection. It is illustrated in Figure 28.11 for a predicate focus construction in English. There are three main components of this projection. Basic information units correspond to the information content captured by a simple WH-word like *who*, *what*, or *where*. In simple sentences this notion may seem redundant with syntactic phrases, but it plays an important role in the analysis of information structure in complex sentences. The second component is the actual focus domain, which is what is actually in focus in a given context; the elements in small caps in (9) are in the actual focus domain in those examples. The third component, which was introduced in RRG and is not part of Lambrecht's original account, is the potential focus domain. Languages differ as to constraints on where the actual focus domain can be in a clause. In some like English, it can fall on

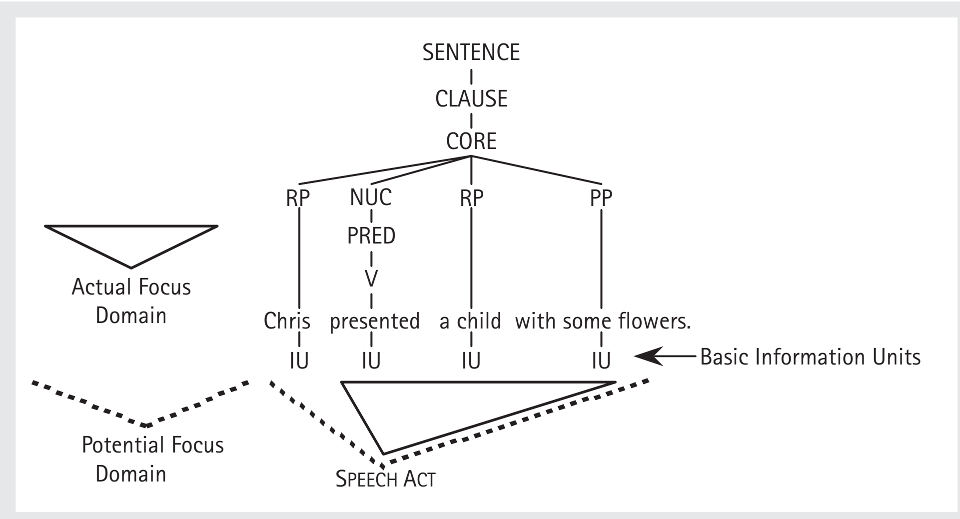


Figure 28.11. Focus structure projection of an English predicate-focus construction

any word or phrase, as (9) shows. In others, e.g., Italian, it is excluded from the preverbal core position and can only include the nucleus and what follows (see VVLP §5.4, Van Valin 1999 for detailed discussion). The potential focus domain is a feature of the grammar of the language, while the actual focus domain is contextually determined. In Van Valin (2005) formal representations of context based on Discourse Representation Theory (Kamp and Reyle 1993, von Heusinger 1999) are incorporated into the theory, in order to derive the different focus types. They can also play an important role in linking in some languages (see Van Valin 2005 §5.4.1, Shimojo 2008, 2009). A very new development in the theory is an explicit representation of prosody (O'Connor 2008), which will be incorporated into the focus structure projection.

It is possible to represent all three projections in a single tree, as in Figure 28.12. It should be noted that these are not three separate representations of the sentence; rather, they are representations of three types of information which are simultaneously present in the sentence.

28.2.4 Grammatical relations

As noted in §28.1, in the earliest work on RRG it was argued that grammatical relations like subject and direct object are not universal and cannot be taken as the basis for adequate grammatical theories. In place of these notions, RRG employs the notion of "privileged syntactic argument" [PSA], which is a construction-specific relation and is defined as a restricted neutralization of semantic roles and pragmatic functions for syntactic purposes. The other arguments in a clause are characterized

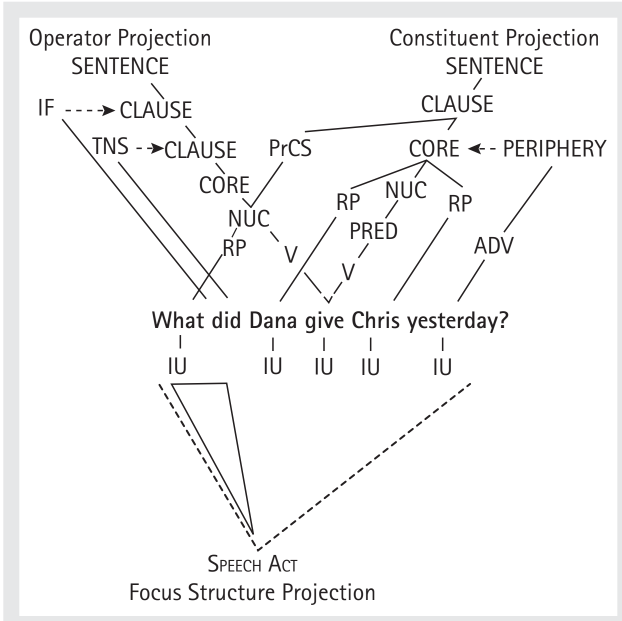


Figure 28.12. English sentence with all three projections represented

as direct or oblique core arguments; there is nothing in RRG corresponding to direct or indirect object (see Van Valin 2005, chapter 4).

Languages have selection hierarchies to determine the PSA; the two main ones are given in (11).

(10) Privileged syntactic argument selection hierarchy:
Arg of DO > 1st arg of **do**′ > 1st arg of **pred**′ (x,y) > 2nd arg of **pred**′ (x,y) > **pred**′ (x)

(11) Privileged Syntactic Argument Selection Principles
a. Accusative construction: Highest ranking direct core argument in terms of (10)-default
b. Ergative constructions: Lowest ranking direct core argument in terms of (10)-default
c. Restrictions on PSA in terms of macrorole status:
1. Languages in which only macrorole arguments can be PSA: German, Italian, Dyirbal, Jakaltek, Sama, etc.
2. Languages in which non-macrorole direct core arguments can be PSA: Icelandic, Georgian, Japanese, Korean, Kinyarwanda, etc.

The PSA selection hierarchy in (10) is the actor part of the AUH. For a language like English, (11a) captures the fact that, in an active voice clause with a transitive verb, the actor is the PSA, whereas for a language like Dyirbal, in an active voice clause with a transitive verb the undergoer is the PSA, following (11b). These are the default choices; it is possible for an undergoer to serve as PSA in a passive construction in an accusative language like English or German, and it is likewise possible for an actor to serve as PSA in an antipassive construction in syntactically ergative languages like Dyirbal and Sama (Philippines; Walton 1986). Languages also differ with respect to whether the PSA must be a macrorole: German, Italian, Dyirbal, Jakaltak (Mayan), and Sama restrict PSA selection to actors and undergoers only, while Icelandic, Georgian, Japanese, and Kinyarwanda allow non-macrorole direct core arguments to function as PSA (see VVLP §7.3.1.1; Van Valin 2005 §4.2).

An aspect of (11a) with significant typological consequences is whether it is a default rule, as in English, German, and many other languages, or whether it is an absolute rule, as in Lakhota, Warlpiri, and many other languages. That is, in a language like Lakhota the highest-ranking argument in the LS is always the PSA; there is no other choice, as the language lacks a voice opposition. With a transitive or other multi-argument verb, the speaker has no choice as to which argument serves as the PSA. The contrast between PSAs in English-type languages and Lakhota-type languages can be captured in a distinction between "variable PSAs" [most English constructions] vs. "invariable PSAs" [all Lakhota constructions].[8] In languages with variable PSAs, in particular constructions, e.g., those like (12) below, one of the factors affecting which argument may be selected as PSA is information structure. It has long been recognized that there is a strong tendency for the RP selected as PSA to be the most topical in the particular context. In RRG variable PSAs in such constructions are analyzed as "pragmatically-influenced PSAs"; it should be noted that not all variable PSAa are pragmatically influenced. There is no pragmatic influence on PSA selection with invariable PSAs.

Two typological points need to be mentioned. First, invariable ergative PSAs, i.e., (11b), are extremely rare but not non-existent. Second, the most common type of PSA cross-linguistically seems to be invariable accusative PSAs. VVLP §6.5 presents extensive arguments for both of these claims and proposes an explanation for these facts.

PSAs may be characterized functionally as controllers or pivots. These two functions are exemplified in (12) and (13).

(12) a. The tall man$_i$ hit William$_j$ and then ___$_{i/*j}$ ran away.
CONTROLLER PIVOT
b. William$_j$ was hit by the tall man$_i$ and then ___$_{*i/j}$ ran away.
CONTROLLER PIVOT

(13) a. Bill$_i$ persuaded the tall man$_j$ [___$_{*i/j}$ to visit Leslie].
CONTROLLER PIVOT
b. The tall man$_j$ was persuaded by Bill$_i$ [___$_{*i/j}$ to visit Leslie].
CONTROLLER PIVOT

[8] While it is normally the case that languages with variable PSAs have a voice opposition, it is not always the case. See Van Valin (2009) for discussion of two languages, Liangshan Nuoso and Barai, which have variable PSAs but lack a formal voice opposition.

Pivots are canonically the missing argument in a construction, as in (12) and (13), while controllers prototypically supply the interpretation for a pivot. It should be noted that there can be pivots without controllers, e.g., the extracted element in an extraction construction, and controllers without pivots, e.g., reflexive controllers. A further contrast is highlighted in these examples, the contrast between syntactic and semantic pivots and controllers. In the construction in (12), the controller is the first RP in the core, the traditional "subject", regardless of its semantic function, whereas in the construction in (13), the controller is the undergoer argument, regardless of its syntactic status. Hence the controller in (12) is a syntactic controller, while the controller in (13) is a semantic controller. The types of pivots and controllers that the constructions of a language have are typologically very significant.

The RRG position on the universality of grammatical relations can be summarized as follows. For a language to have grammatical relations in the usual sense of syntactic relations which are not reducible to semantic roles, it must have at least one construction with a syntactic pivot and/or a syntactic controller. There are languages, e.g., Acehnese (Durie 1985, 1987), which have only semantic pivots and controllers, and therefore they lack grammatical relations in the purely syntactic sense. Furthermore, in languages with syntactic pivots and controllers, there is variation in terms of whether they pattern accusatively, ergatively, or some other pattern, and whether they have variable and invariable PSAs or only invariable PSAs. The grammatical relations in the former type of language are not the same as those in the latter. Hence even among the vast majority of languages with syntactic pivots and controllers, there is no uniformity as to the nature of the PSAs.

28.2.5 The structure of complex sentences

The three central components of the LSC also turn out to be the three fundamental building blocks of complex sentences in human language. The unmarked pattern for the construction of complex sentences involves combining nuclei with nuclei, cores with cores, clauses with clauses, or sentences with sentences. These are called levels of "juncture" in RRG, i.e., nuclear juncture, core juncture, clausal juncture, and sentential juncture. Sentential junctures are complex constructions made up of multiple sentences, while clausal junctures involve sentences containing multiple clauses. Examples of nuclear junctures from French, English, and Mandarin are given in (14) and the representation of (14a) is in Figure 28.13. Justifications for these structures can be found in Van Valin (2005).

(14) a. Je ferai manger les gâteaux à Jean.
1sg make.FUT eat the cakes to John
'I will make John eat the cakes.'
[two nuclei, *faire* and *manger*, in a single core]

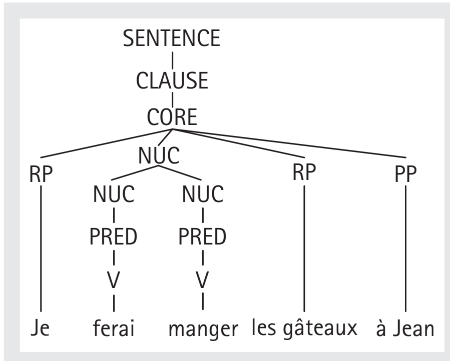


Figure 28.13. The structure of (14a)

b. John forced open the door.
[two nuclei, *force* and *open*, in a single core]

c. Tā qiāo pò le yí ge fànwǎn.
3sg hit 'break' PRFV one CL bowl
'He broke (by hitting) a ricebowl.'
[two nuclei, *qiāo* 'hit' and *pò* 'break', in a single core] (Hansell 1993)

Core junctures involve two or more cores (which may themselves be internally complex) in a clause. Examples from French, English, and Mandarin are given in (15), and the structure of (15a) is presented in Figure 28.14. In this type of core juncture, the two cores share a core argument; "sharing a core argument" is defined formally in terms of the linking algorithm mapping syntactic and semantic representations into each other.

(15) a. Je laisserai Jean manger les gâteaux.
1sg let.FUT John eat the cakes
'I will let John eat the cakes.'

b. I ordered Fred to force the door open.

c. Tā jiāo wǒ xǐe zì.
3sg teach 1sg write characters
'She teaches me to write characters.'

Of equal importance in the RRG theory of complex sentences is the set of possible syntactic and semantic relations between the units in a juncture; the semantic relations are discussed below. The syntactic relations between units are called "nexus" relations in RRG. Traditionally, only two basic nexus relations are recognized, coordination and subordination. Subordination is divided into two subtypes, daughter subordination and peripheral subordination. They are illustrated in Figure 28.15.

The embedded clause in the first sentence is a daughter of the core node, while in the second the embedded clause is an adjunct in the periphery modifying the core.

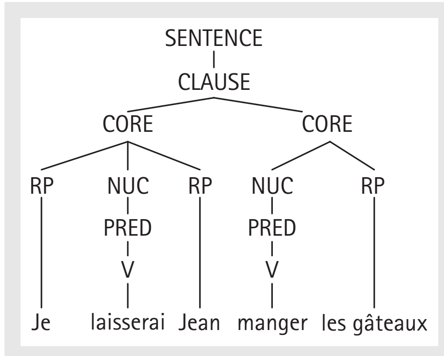


Figure 28.14. The structure of (15a)

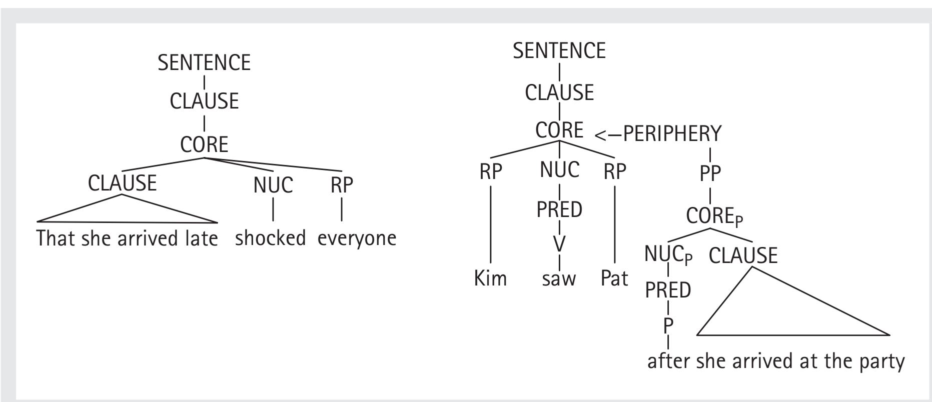


Figure 28.15. Daughter and peripheral subordination at the core level in English

In addition to distinguishing two types of subordination, RRG, following Olson's (1981) analysis of clause linkage in Barai (a Papuan language), posits a third nexus type: "cosubordination", which is essentially tight, dependent coordination. The dependence is operator dependence; that is, in cosubordination, the units obligatorily share one or more operators at the level of juncture. In the Mandarin example in (14c), aspect obligatorily has scope over both nuclei, and therefore the nexus is cosubordination. This is represented as in Figure 28.16.

The following examples from Turkish (Watters 1993) illustrate obligatory operator sharing and the lack of it in Turkish core cosubordination and coordination, respectively. The term "coordination" here is being used for an abstract linkage relation referring to a relationship of equivalence and operator independence at the level of juncture. It is distinct from conjunction, which is a construction type of the general form "X conj Y", which may be one of the formal instantiations of coordinate nexus.

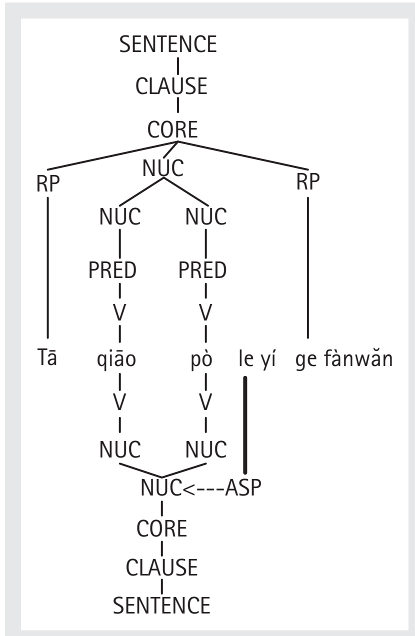


Figure 28.16. Nuclear cosubordination in Mandarin

(16) a. Core cosubordination
Gid-ip gör-meli-yiz.
go-CMPL see-MODAL-1pl
'We ought to go and see.'

b. Core coordination
Müzik dinle-yerek, uyu-yabil-ir-im.
music listen-CMPL sleep-MODAL-AOR-1sg
'While listening to music, I can sleep.'
(Not, 'while I am able to listen to music, I am able to sleep.')

In (16a), the modal operator -*meli*- "ought" has scope over both cores, and therefore the nexus is cosubordinate; in (16b), on the other hand, the modal operator -*yabil*- "able" has scope only over the final core, hence coordinate nexus. The structural representations for (16a,b) are given in Figure 28.17.

The following sentences from Kewa (Franklin 1971) are a minimal triple for the three nexus types at the clause level.

(17) a. Nipú ípu-la pare ní paalá na-pía. Coordination
3sg come-3sgPRES but 1sg afraid NEG-be.1sgPRES
'He is coming, but I am not afraid.'

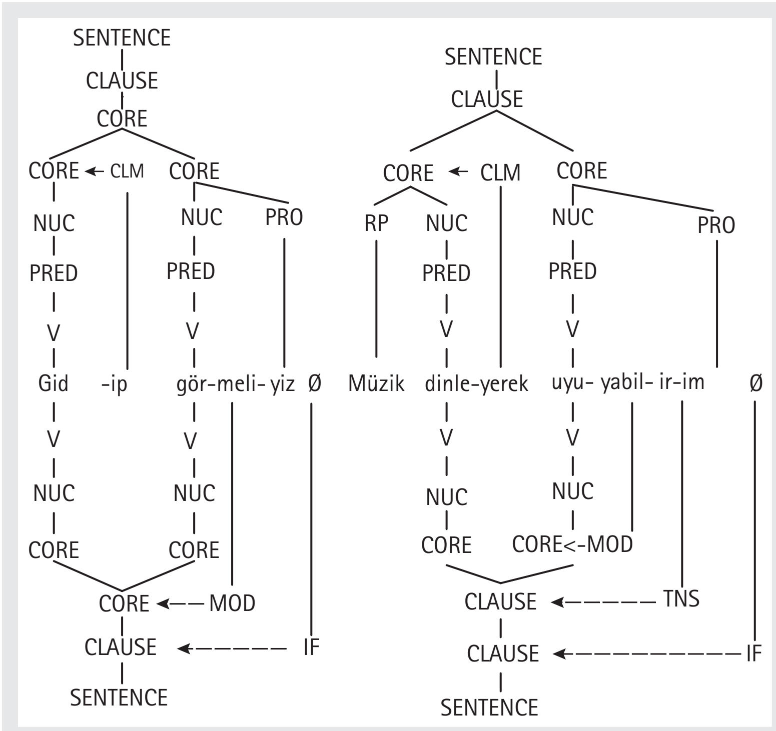


Figure 28.17. Turkish Core Junctures

b.	(Ní)	Épo	lá-ri	épa-wa.	Cosubordination
	(1sg)	whistle	say-SIM.SS	come-1sgPAST	

‘I whistled while I came,’ or ‘I came whistling.’

c.	(Ní)	Épo	lá-lo-pulu	irikai	épa-lia.
	(1sg)	whistle	say-1sgPRES-CAUSAL	dog	come-3sgFUT

‘Because I am whistling, the dog will come.’ Subordination (peripheral)

The four levels of juncture combine with the three nexus types to generate eleven possible complex sentence types; there is no sentential cosubordination because there are no sentence-level operators, hence no possible operator sharing. In addition, both subtypes of subordination are possible at the clause, core, and nuclear levels. Not all of them are instantiated in every language. The juncture-nexus types found in a language may be realized by more than one formal construction type; for example, both *Mary sat playing the guitar* and *Robin tried to open the door* instantiate core cosubordination, while both *For Sam to leave now would be a mistake* and *Lisa's losing her job shocked everyone* instantiate core subordination in English. The juncture-nexus types may be ordered into a hierarchy in terms of the tightness of the syntactic link between the units, i.e., in terms of how integrated the units are into a single unit or are coded as distinct units. This is given in Figure 28.18.

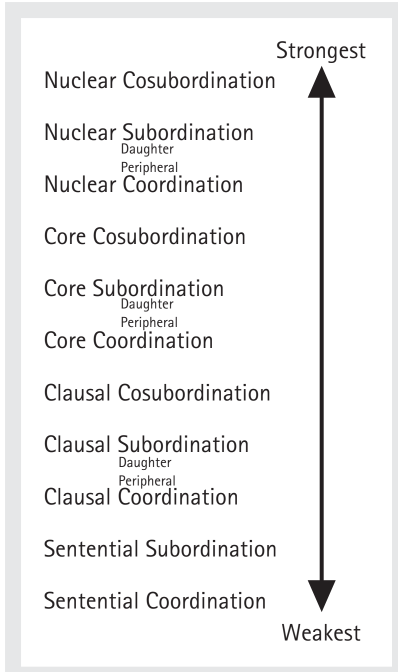


Figure 28.18. Interclausal syntactic relations hierarchy

The syntactic clause-linkage relations discussed earlier are used to express certain semantic relations between the units in the linkage, e.g., causation, purpose, and temporal sequence. The interclausal semantic relations are given in (18).

(18) Interclausal Semantic Relations

a. Causative [1]: the bringing about of one state of affairs directly by another state of affairs, usually an event or action, e.g., *Max painted the door green, Larry pushed the door open*.
b. Phase: a separate verb describes a facet of the temporal envelope of a state of affairs, specifically its onset, its termination, or its continuation, e.g., *Chris started crying, Fred kept singing, Hari finished writing the chapter*.
c. Modifying subevents
    1. Manner: the manner in which a motion event is carried out, e.g., *Bill entered the room skipping*.
    2. Motion: motion accompanying another action, e.g., Mparntwe Arrerente *angk-tyantye-*[speak-go.upward] 'speak while going up' (Wilkins 1991).
    3. Position: stance while doing an action, e.g., *Dana sat reading a newspaper*.

4. Means: the means by which an action is carried out, e.g., *Sam opened the box by slicing it with a knife.*

d. Psych-action: a mental disposition regarding a possible action on the part of a participant in the state of affairs, e.g., *Max decided to leave, Sally forgot to open the window, Tanisha wants to go to the movies.*
e. Purposive: one action is done with the intent of realizing another state of affairs, e.g., *Juan went to the store to buy milk, Susan brought the book to read.*
f. Jussive: the expression of a command, request, or demand, e.g., *Pat asked the student to leave, The king ordered the troops to attack the city.*
g. Causative [2]: the bringing about of one state of affairs through a distinct action or event, e.g., *Fred forced Max to paint the table.*
h. Direct perception: an unmediated apprehension of some act, event, or situation through the senses, e.g., *Rex saw the child open the door, Yolanda heard the guests arrive.*
i. Indirect perception: the deduction of some act, event, or situation from evidence of it, e.g., (looking at an empty desk)*I see that John has gone home early.*
j. Propositional attitude: the expression of a participant's attitude, judgment, or opinion regarding a state of affairs, e.g., *Carl believes that UFOs are a menace to the earth, Paul considers Carl to be a fool, Most fans want very much for their team to win.*
k. Cognition: an expression of knowledge or mental activity, e.g., *Aaron knows that the earth is round, George is thinking about Madeleine's refusal to go out with him.*
l. Indirect discourse: an expression of reported speech, e.g., *Frank said that his friends were corrupt.*
m. Direct discourse: the direct quotation of a speech event, e.g., *Frank said, "My friends are corrupt."*
n. Circumstances: the spatial or temporal parameters of an event, e.g., *Sam talked to Sally at the library after work.*
o. Reason: the motivation or cause for an action or event, e.g., *The baby cried because she was hungry.*
p. Conditional: an expression of what consequence would hold, given the conditions in a particular state of affairs, e.g., *If it rains, we won't be able to have a picnic, Were Fred to leave now, he would look like a fool.*
q. Concessive: the content of the main clause holds unexpectedly, given the content of the subordinate clause, e.g., *Bill made it to work, even though it was snowing heavily.*
r. Temporal
   1. Simultaneous states of affairs: one state of affairs is temporally cotermi nous with another, e.g., *Max danced and Susan played the piano, Kim had chicken pox and at the same time Leslie had the measles.*

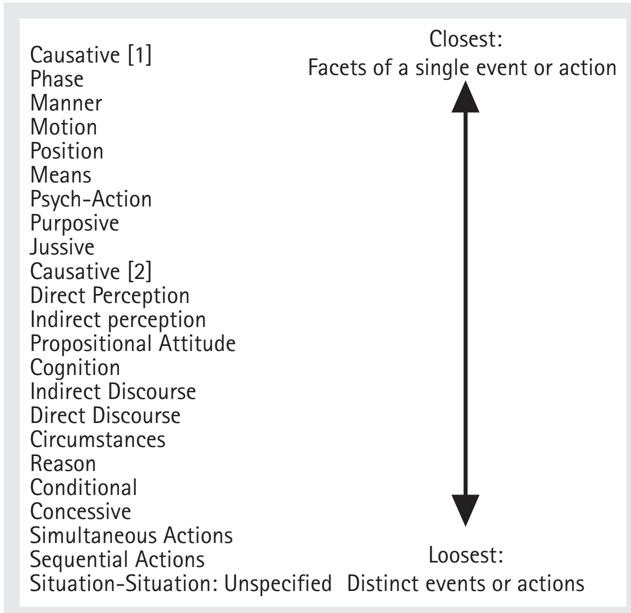


Figure 28.19. Interclausal semantic relations hierarchy

2. Sequential states of affairs: one state of affairs follows another temporally, with or without any temporal overlap, e.g., *Juan had finished talking, and then Carlos entered the room, Vidhu was sitting down, and the band began to play*.

s. Temporally unordered states of affairs: the temporal relation between states of affairs is unexpressed, e.g., *Tyrone talked to Tanisha, and Yolanda chatted with Kareem*.

These relations may be formalized in terms of the same decomposition used for verbs (see Van Valin 2005: 207–8, also Ohori 2001, 2005).

The semantic relations form a continuum expressing the degree of semantic cohesion between the propositional units linked in the complex structure, i.e., the degree to which they express facets of a single action or event or discrete actions or events. This may be represented as in Figure 28.19.

The syntactic linkage relations are ranked hierarchically in terms of the strength of the syntactic bond between the units in Figure 28.18. The interaction of the two hierarchies is expressed in the interclausal relations hierarchy in Figure 28.20. The relationship between the syntactic and semantic relations in clause linkage is very complex; i.e., it is not one-to-one, but there are some striking regularities cross-linguistically. The primary principle governing the interaction of the two hierarchies is iconic: the closer the semantic relation between two propositions is, the stronger the syntactic link joining them (Silverstein 1976, Givón 1980). In other words, the semantic relations at the top end of the hierarchy should be realized

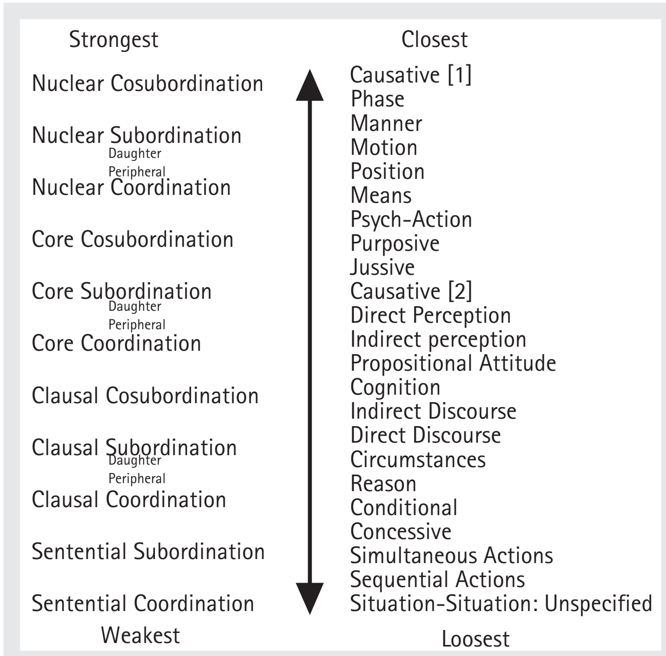


Figure 28.20. Interclausal relations hierarchy

by the linkage categories at the top as well, and the relations at the bottom of the hierarchy should be realized by the linkage categories at the bottom of the syntactic side. Moreover, while there is often more than one syntactic realization of a particular semantic relation, the tightest syntactic linkage realizing it should be tighter than the tightest syntactic linkage realizing looser semantic relations.

28.3 Linking between syntax and semantics

All of the components of the RRG linking system have been introduced. This is summarized in Figure 28.21.

Logical structures, macroroles, and the hierarchy linking them are universal in that there is very little cross-linguistic variation; this is the domain of lexical processes. Where languages differ substantially is how macroroles and other arguments link into the syntax.

The reason the arrows in Figure 28.21 are double-headed is that the linking system works both from semantics to syntax and from syntax to semantics. A theory which

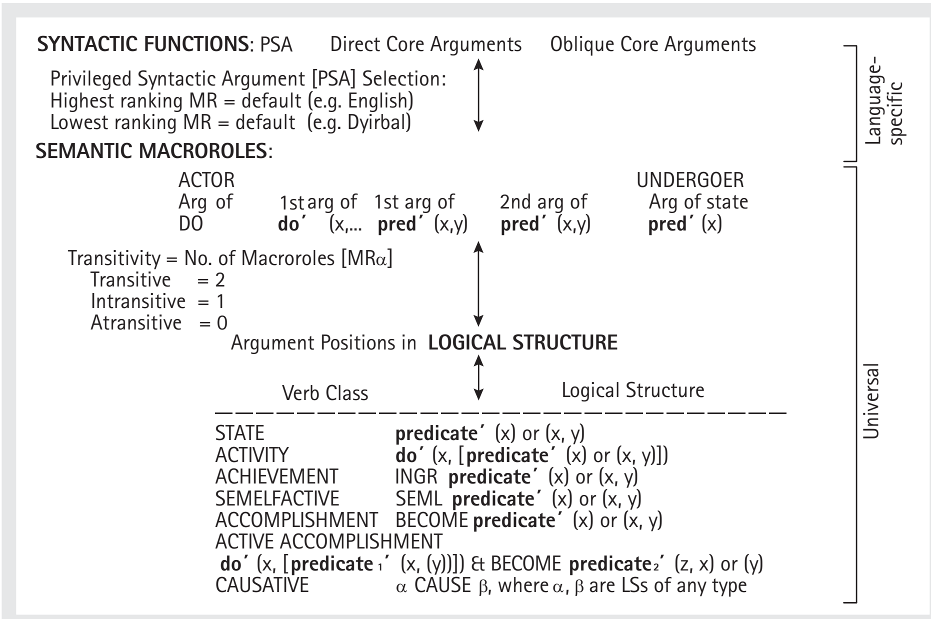


Figure 28.21. RRG Linking System

could describe the linking from semantics to syntax only could be part of a language production system, but it would not be adequate for a comprehension system. In such a system, the parser, as an idealization, would take the input and produce a structured syntactic representation of it, identifying the elements of the layered structure of the clause and the cases, adpositions and other grammatically relevant elements in the sentence. It is then the grammar's job to map this structure into a semantic representation, as the first step in interpreting it, and this is where the syntax to semantics linking algorithm is required. The details of the linking algorithms are given in Van Valin (2005).

The linking between syntax and semantics is governed by a very general principle called the "Completeness Constraint"; it states simply that all of the specified arguments in the semantic representation of a sentence must be realized in the syntax in some way, and conversely that all of the expressions in the syntax must be linked to something in the semantic representation of a sentence, in order to be interpreted.

An important part of the linking involves finite verb agreement, case assignment, and preposition assignment. The finite verb agreement rule for accusative languages like English, German and Croatian is given in (19).

(19) Finite verb agreement in Croatian, German, and Icelandic:
The controller of finite verb agreement is the highest-ranking core macrorole argument (in terms of (10)).

The rule is not formulated with respect to any syntactic position or function, or with respect to any case. Case assignment rules are formulated in a similar way. The basic rules for direct core arguments in accusative languages are given in (20) and for ergative languages in (21); these do not pertain to case assigned by adpositions.

(20) Case marking rules for accusative languages:
  a. Highest-ranking core macrorole (in terms of (10)) takes nominative case.
  b. Other core macrorole takes accusative case.

(21) Case marking rules for ergative languages:
  a. Lowest-ranking core macrorole (in terms of (10)) takes absolutive case.
  b. Other core macrorole takes ergative case.

In addition, there is a rule for dative case assignment, which applies to both systems.

(22) Assign dative case to non-macrorole direct core arguments (default).

Dative case is assigned only when the rules for the other cases cannot apply.[9] In a language like English without RP case marking, there are rules for preposition assignment (Jolly 1993). The rules for *to* and *from* are given in (23).[10]

(23) Preposition assignment rules for English
  a. Assign *to* to NMR *x* argument in LS segment: …BECOME/INGR **pred**′ (x, y)
  b. Assign *from* to NMR *x* argument in LS segment: …BECOME/INGR NOT **pred**′ (x, y)

The rule in (23a) is particular important for the "dative shift" verbs in English, e.g., *give*, *send*, *show*, etc. The alternation in (24) is handled in terms of variable undergoer selection; both sentences would have the same LS, given in (24c).

(24) a. Mary sent a letter to Sally.
  b. Mary sent Sally a letter.
  c. [**do**′ (Mary, Ø)] CAUSE [BECOME **have**′ (Sally, letter)]

In (24a) undergoer selection reflects the default choice in terms of the AUH in Figure 28.10, i.e., the rightmost argument in LS is chosen to function as undergoer. In (24b), on the other hand, the second lowest-ranking argument (which is also the second highest-ranking), *Sally*, is selected as undergoer. In (24a), the conditions for (23a) are met, and therefore Sally is marked by *to*. In (24b), however, this is not met, and therefore it does not apply. Alternations with verbs like English *present*, German *schenken* vs. *beschenken* "give as a gift", Croatian *darovati* "give as a gift", and Dyirbal *wugal* "give" are all analyzed as instances of variable undergoer selection (see Van Valin 2005: §4.4, Van Valin 2007).

[9] There is also a rule for assigning instrumental case, but it is complex and not necessary for this discussion. See Van Valin (2005), §4.4.

[10] There is also a rule for assigning *with*, which is very similar to the rule for instrumental case.

Most of what counts as "syntax" in many theories is handled in RRG in terms of constraints on the semantic representation, in terms of information structure, or in syntactic phase of the linking (see the analysis of WH-questions below). The analysis of reflexivization in RRG follows the approach in Jackendoff (1992) and states the hierarchical constraints for core-internal ("clause-bound" in other theories) reflexivization at the LS level, not with respect to the syntactic representation. The principles affecting the scope and interpretation of quantifiers are related to information-structure contrasts, not phrase structure. RRG treats constructions as an important part of syntax, and they are represented in terms of constructional schemas. Cross-constructional and cross-linguistic generalizations are captured in terms of the general principles and constraints that constitute the linking algorithms, e.g., the actor–undergoer hierarchy, the layered structure of the clause, the PSA selection hierarchy. Only the idiosyncratic, language-specific features of constructions are represented in constructional schemas, which may include syntactic, morphological, semantic, and pragmatic (focus structure) information.

A simple example from English illustrating the operation of the semantics-to-syntax linking algorithm is given in Figure 28.22. The numbers refer to the general steps of the algorithm: (1) constructing the semantic representation of the sentence in the lexicon; (2) assigning actor and undergoer; (3) determining PSA selection, case and adposition assignment, and agreement; (4) selecting the appropriate syntactic template from the syntactic inventory; and (5) linking the elements from the semantic representation into the appropriate positions in the syntactic representation. The numbers in the diagram, especially 2–3, should not be interpreted as indicating steps in a derivation. Rather, they signal steps in the linking which involve adding morphosyntactic information to the semantic representation. The output of step 3 could equally well be represented as "… [**do**′ (Sandy[Actor, *by*-ACC], Ø)] CAUSE [BECOME **have**′[passive, 3pl] (Chris[NMR, *to*-ACC], flowers[Und, PSA, NOM])] …".

Because this sentence is a passive, the undergoer appears as the "subject", with the actor appearing in a peripheral PP marked with *by*. These language-specific details would be represented in the constructional schema for the English passive, given in Table 28.4.

The information in the constructional schema is a combination of general principles (template selection principles, PSA selection principles, general characterization of non-default PSA selection), plus language-specific information, e.g., the form of the verb and the choice of auxiliary. See Van Valin (2005) for detailed discussion and explication of all of these points.

A simple example of the linking from syntax to semantics is given in Figure 28.23. Here again the numbers refer to the general steps in the algorithm: (1) extract all of the information possible from the overt morphosyntactic form of the sentence, including the voice of the verb (if the language has voice), case marking, word order, and adpositions; (2) retrieve the LS of the predicate in the nucleus from the lexicon

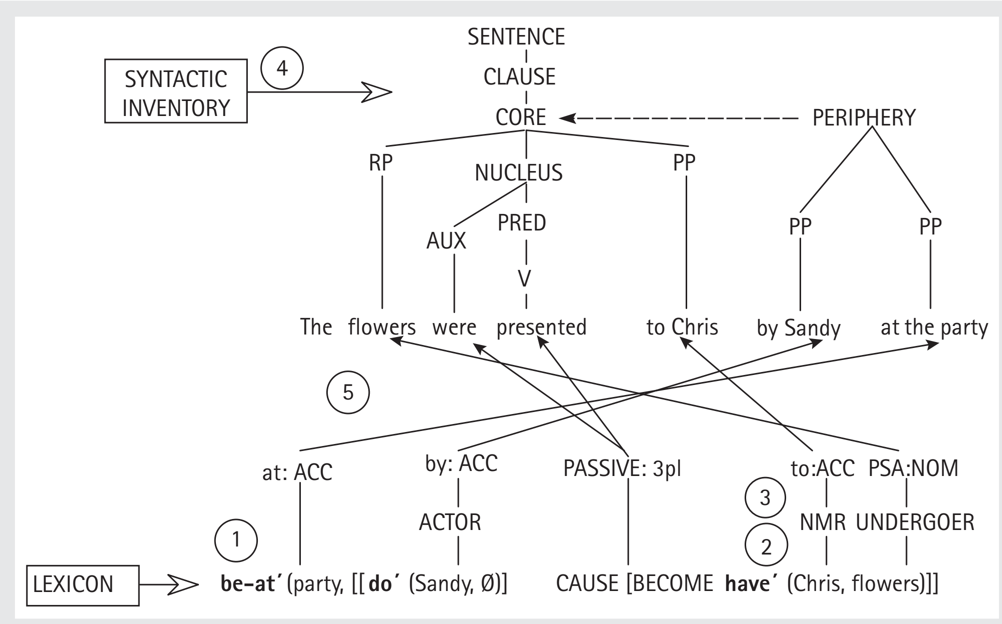


Figure 28.22. Linking from semantics to syntax in a simple sentence in English

Table 28.4. Constructional schema for English passive (plain)

CONSTRUCTION: English passive (plain)
SYNTAX: Template(s): (following template selection principles; not given above) PSA: (11a,c2), Variable [± pragmatic influence] Linking: Undergoer to PSA; Actor omitted or in peripheral *by*-PP MORPHOLOGY: Verb: past participle Auxiliary: *be* SEMANTICS: PSA is not instigator of state of affairs but is affected by it (default) PRAGMATICS: Illocutionary force: Unspecified Focus structure: No restrictions; PSA = topic (default)

and assign macroroles to the extent possible; and (3) link of the information derived from steps (1) and (2). The syntactic representation is produced by the parser, which turns the acoustic input into a labeled syntactic representation.

The linking in a WH-question in English, in both directions, is illustrated in Figures 28.24 and 28.25; the linking of the peripheral adjunct *yesterday* is not represented. In the linking from semantics to syntax in Figure 28.24, the undergoer *what*

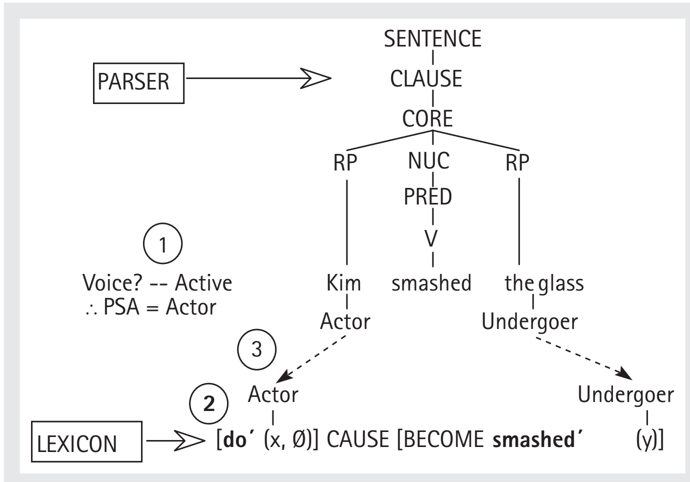


Figure 28.23. Linking from syntax to semantics in a simple sentence in English

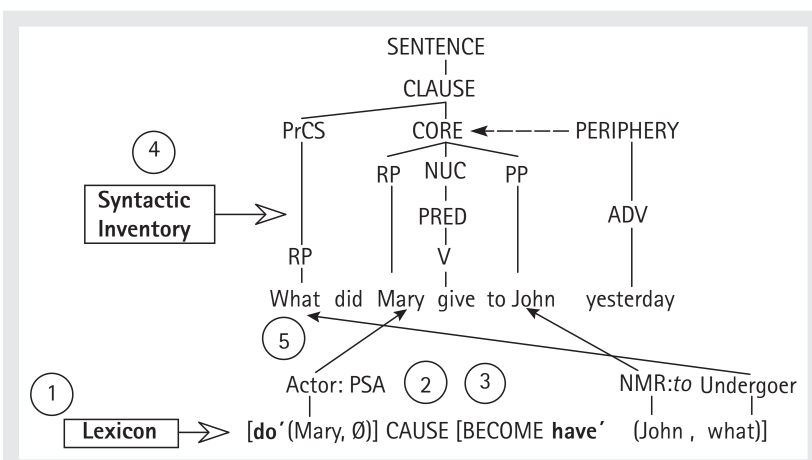


Figure 28.24. Linking from semantics to syntax in a WH-question in English

is linked directly to the PrCS; there is no empty argument position in the core, i.e., no trace. The rule in (23a) applies to assign *John* the preposition *to*.

There are two important complications in the syntax to semantics linking with this sentence, as shown in Figure 28.25. First, no conclusion can be drawn from the morphosyntax regarding the function of *what*; hence in step 1 this is simply labeled "RP". Second, because *give* is a variable undergoer selection verb, it is not possible to assign undergoer to an argument in the LS in step 2, unlike in Figure 28.23. So the linking of *John* is determined by using the inverse of (23a) as a linking

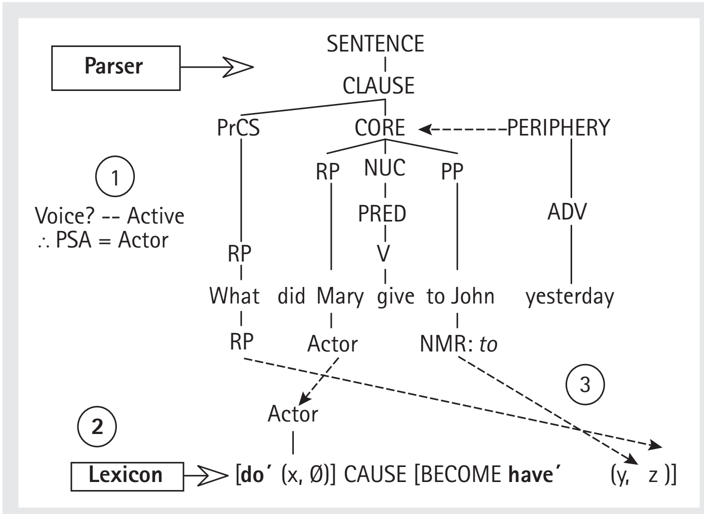


Figure 28.25. Linking from syntax to semantics in a WH-question in English

principle: since *John* is marked by *to*, it must be the first argument of **have**′. After *Mary* is linked to the *x* argument and *John* to *y*, the Completeness Constraint forces the linking of *what* to the *z* argument, which yields the correct interpretation. Constraints on WH-question formation and other "extraction" constructions are explained in terms of the interaction of information structure and syntax, in particular in terms of restrictions on the potential focus domain (Van Valin 1995, 1998, 2005).

28.4 Conclusion

The more complete picture of RRG that emerges from this discussion is given in Figure 28.26.

In the linking from semantics to syntax, the source of the syntactic representation is the templates of the syntactic inventory. In the syntax to semantics linking, the source is the parser. The lexicon plays an important role in both. Discourse-pragmatics, i.e., information structure, interacts with the linking algorithm in significant ways at various steps. Constructional schemas provide the language-specific morphosyntactic information that complements the general, cross-linguistically valid principles of the theory.

In §28.1, the motivating questions for RRG were presented: "what would a linguistic theory look like if it were based on the analysis of languages with diverse

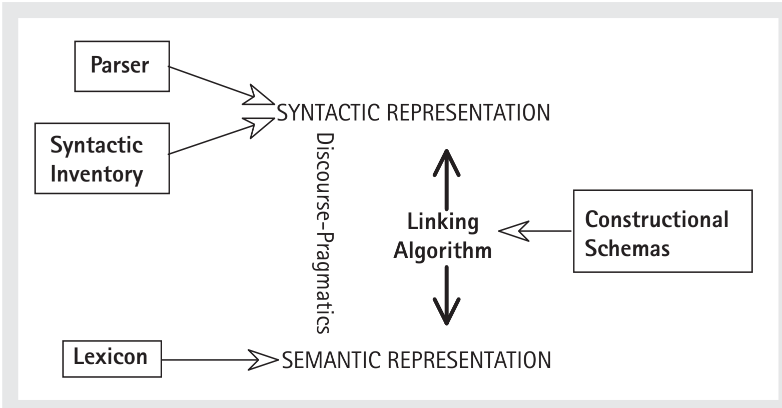


Figure 28.26. Organization of Role and Reference Grammar (final)

structures, such as Lakhota, Tagalog, Dyirbal, and Barai (Papua New Guinea), rather than on the analysis of English?", and "how can the interaction of syntax, semantics, and pragmatics in different grammatical systems best be captured and explained?" This chapter has sketched out the answer to these questions, with emphasis on the first. Because of this typologically diverse starting point, the constructs presented herein differ in significant ways from those in other theories: a non-endocentric theory of phrase and clause structure, the concept of reference phrase instead of noun phrase, rejection of the universality of grammatical relations, the construction-specific notion of PSA, cosubordination as a third linkage type in addition to coordination and subordination, etc. These typologically motivated theoretical and descriptive constructs have made RRG a useful framework for linguists primarily concerned with language description, and yet it makes possible the analysis and explanation of the kind of morphosyntactic phenomena which have been the focus of much of the theoretical work of the past decades.

CHAPTER 29

THE ANALYSIS OF SIGNED LANGUAGES

SHERMAN WILCOX
PHYLLIS PERRIN WILCOX

29.1 What are signed languages?

Signed languages[1] are natural human languages used by deaf people throughout the world as their native or primary language. Although no formal survey of the world's signed languages has ever been conducted, it is generally recognized by linguists that they number in the hundreds; the 13th edition of the Summer Institute of Linguistics *Ethnologue* of the world's languages lists 103 signed languages (Grimes 1996).

Like spoken languages, signed languages may be classified into genetic or family groups. These genetic relations follow the historical development of signed languages, and so do not reflect the same relations that may exist for spoken languages in the same areas. For example, French Sign Language is a parent language of both American Sign Language (ASL) and Russian Sign Language.

[1] Many authors refer to "sign languages". We prefer the term "signed language", parallel to spoken language and written language.

29.1.1 Analyzing visual-gestural languages

The linguistic analysis of signed language form was at the heart of the controversy over whether signed languages are natural languages. At the root of this controversy was the question of duality of patterning: do signed languages exhibit duality of patterning, or are the words (the "signs") of signed languages holistic units, depictive gestures consisting of no sublexical or submorphemic internal structure?

Modern linguistic analysis of ASL began with the pioneering work of William C. Stokoe. In *Sign Language Structure*, Stokoe (1960) offered the first linguistic analysis of ASL form, demonstrating that it exhibits duality of patterning. Stokoe analyzed signs into three major phonological classes: handshape (the configuration that the hand makes when producing the sign), location (the place where the sign is produced), and movement (the movement made in producing the sign). He termed these meaningless units of formation cheremes, the signed equivalent of phonemes in spoken languages. Battison (1978) later proposed adding a fourth major class, orientation (the direction in which the palm faces when producing the sign). While Stokoe realized that the articulation of certain signs exhibited internal sequentiality and recorded this fact in his notational system, he believed that the major units of organization, the cheremes, are simultaneously rather than sequentially organized. In order to grasp this concept, Stokoe (1980: 369) wrote that one must look at signs "in different ways if different *aspects* of their structure are to be seen":

> In producing a sign language utterance, some part (or parts) of the signer's body acts. If the active part is mobile enough, there are various places in which the action may occur, i.e., begin, take place, or end. But the action, the active part, and the place are all present simultaneously. The problem is to see what composes a sign (i.e., what elements they can be decomposed into) when signs are taken as equivalents of words or morphemes of spoken languages. Signs cannot be performed one aspect at a time, as speakers can utter one segment of sound at a time. Signers can of course display handshapes of manual signs ad libitum, but they cannot demonstrate any significant sign action without using something to make that action somewhere. By an act of imagination, however, it is possible to "look at" a sign *as if* one could see its action only or its active element only or its location only. In this way *three aspects* of a manual sign of sign language are distinguished, not by segmentation, it must be reemphasized, but by imagination. (Stokoe 1980: 369)

Signed language phonology advanced rapidly after Stokoe's initial insight that signs have parts. Linguists began applying existing phonological models such as autosegmental (Sandler 1986), prosodic (Brentari 1998), and moraic (Perlmutter 1989) to signed language. One phonological model that has received widespread attention is the movement-hold model first proposed by Liddell (1984). This model argued against Stokoe's original view of a sign as a simultaneous bundle of primes. Liddell proposed instead that the majority of signs are segmentable into sequentially-organized movements and holds. In this scheme, movements and holds become the segmental units. Each is described by simultaneously-organized feature bundles. Thus, in the movement-hold model, signed words/morphemes and spoken words/morphemes are identically structured.

Stokoe returned to the formational structure of signed language in an essay in which he introduced the term "semantic phonology". In this article, Stokoe sought to simplify what he regarded as overly complex models of sign phonology, including his own original three-part conception of handshape, location, and movement:

> What I propose is not complicated at all; it is dead simple to begin with. I call it semantical phonology. It invites one to look at a sign . . . as simply a marriage of a noun and a verb . . . [O]ne needs only to think of a sign as something that acts together with its action.
>
> (Stokoe 1991: 107)

The term "semantic phonology" was meant to unify two facts about signed language. First, Stokoe recognized that the phonological primes of a signed language, hands and their shapes, locations, and movements—or, in his simplified view, *the thing that acts with its action*—have inherent semantic or conceptual import in the same way that cognitive linguists attribute conceptual import to grammatical categories such as noun and verb (Langacker 1987*b*). Second, Stokoe believed that this archetypal conceptual structure, in its most basic form *something that acts*, corresponded to the nature of human vision and the difference of retinal cells receptive to detail ("something") and movement ("acts").

Semantic phonology was also a reflection of Stokoe's life-long interest in the evolution of language and his support of a gestural model of language origins (Stokoe 1974; Armstrong et al. 1995). Seen in this light, semantic phonology was more akin to models describing the neurology of actions than to formal linguistic notions of semantics and phonology. The predictions it makes are not so much about phonological rules as it is about the precursors to human language, such as the proposal by Rizzolatti and Arbib (1998: 193) that "language in humans . . . evolved from a basic mechanism originally not related to communication: the capacity to recognize actions".

29.2 Typology of signed languages

Research on the typology of signed languages is still in its infancy. For many years, linguists compared signed languages to the spoken languages of western Europe, where much of the research on signed languages has been done. As more was learned about the linguistics of signed languages, it became clear that this approach led to serious misconceptions about the typology of signed languages. In recent years, Slobin has been a leading proponent of the need to re-examine signed language typology.

Slobin (2005) proposes that signed languages differ typologically from the western European spoken languages that they are commonly compared with along at least two important dimensions. First, signed languages appear to be predominantly head-marked rather than dependent-marked languages. Comparing a construction such as *He/she looks at me* in ASL and Yucatec, Slobin (2005) finds that both rely on polycomponential verbs to indicate the roles of arguments, rather than the use of pronouns that carry grammatical marking of those roles. In ASL, the construction would consist of a single sign, a V-handshape indicating the verb "to look at" (extended index and middle fingers held horizontally) moving from a third-person locus toward the face of the signer. Further, Slobin notes that because argument roles are marked on the verbs, head-marked languages are often considered pro-drop languages.

The second typological feature noted by Slobin is that many if not most signed languages are topic-prominent rather than subject-prominent (Li and Thompson 1976). ASL discourse, for example, is characterized by a prominence of topic-comment structure. Topics are marked by raised eyebrows and a backward head tilt; often there is a pause between the topic and comment phrases, and the final sign of the topic phrase is held slightly longer. Topic-comment structure has been described for ASL as a kind of sentence type along with others, such as questions, imperatives, and assertions (Baker and Cokely 1980), but the relative frequency with which it appears suggests that it is more basic (Janzen et al. 2000). Marked topic phrases serve to manage backgrounded and foregrounded information in signed discourse. Thus topics can appear as part of any given sentence type.

Zeshan (2004) reports on results from a crosslinguistic study based on data from thirty-five geographically and genetically distinct signed languages. Her study investigated the manual and non-manual marking of basic types of questions. Her data showed that the range of crosslinguistic variation is extensive for the structure of question-word paradigms, while other parameters, such as the use of non-manual expressions in questions, show more similarities across signed languages. We will return to the latter finding in section 29.5.3 and suggest a possible explanation for these crosslinguistic similarities.

29.3 Iconicity, metaphor, and metonymy

The cognitive processes of iconicity, metaphor, and metonymy, while recognized in the early years of research on signed languages, have become the object of renewed interest as linguists working in the cognitive linguistic framework have turned their attention to signed languages.

29.3.1 Iconicity

Signed languages are particularly intriguing for linguists interested in the study of iconicity. Early on, linguists recognized the pervasive iconicity of signed languages. Stokoe et al. (1965) noted metaphorical, metonymic, and iconic aspects of ASL. Mandel (1977) described in rich detail several iconic devices that he argued played a vital role in the grammar of ASL. Following this period of interest in the extent of iconicity, linguists began to document constraints on iconicity, examining the loss of iconicity over time in the lexicon (Frishberg 1975), demonstrating that iconicity does not appear to play a role in language acquisition (Meier 1980), and presenting evidence that iconicity does not aid in the processing of signs (Klima and Bellugi 1979). Klima and Bellugi also argued that grammatical processes in ASL work to diminish lexical iconicity. While acknowledging the two faces of signs—"the iconic face and the encoded, arbitrary face"—they observed:

> Grammatical operations that signs undergo can further submerge iconicity. Thus many signs, while having their roots deeply embedded in mimetic representation, have lost their original transparency as they have been constrained more tightly by the linguistic system.
>
> (Klima and Bellugi 1979: 34)

The example that Klima and Bellugi offer is the morphological marking of intensification on certain statives in ASL, expressed phonologically as an initial hold of the sign's movement followed by sudden, rapid release. When this grammatical marker appears on the ASL sign SLOW, they note, the resulting sign means "very slow". Klima and Bellugi pointed out that the sign VERY-SLOW is made with a faster movement than that used in the sign SLOW, and thus they argued that the form of VERY-SLOW is non-iconic with its meaning: VERY-SLOW is articulated very fast.

Wilcox (2004*a*) presented a model of iconicity, which he called *cognitive iconicity*, based on cognitive grammar (Langacker 1987*b*; 1991). A critical claim of cognitive grammar is that both semantic and phonological structures reside within a language user's conceptual space. Conceptual space is multidimensional, encompassing all of our thought and knowledge, "the multifaceted field of conceptual potential within which thought and conceptualization unfold" (Langacker 1987*b*: 76). In this view, similarities among concepts are regarded as distance relations between structures that reside in conceptual space (Gärdenfors 2000). Certain notions reside close to each other in conceptual space because they possess certain similarities. Other notions reside farther apart in conceptual space, reflecting their dissimilarity.

The central claim of cognitive iconicity is that phonological notions also reside in conceptual space. The phonological pole of symbolic linguistic structures reflects our conceptualization of pronunciations, which range from the specific pronunciation of actual words in all their contextual richness to more schematic conceptions, such as a common phonological shape shared by certain nouns or verbs in a particular language. This is the case for ASL, which contains a set of noun–verb pairs in which the form of the noun consists of a spatially reduced, tightly reduplicated movement, while the verb form uses the same handshape and location with an unrestrained, larger movement (Supalla and Newport 1978). Wilcox argues that these forms exhibit two types of iconicity. First, the forms are often iconic for their referent: SWEEP-WITH-BROOM resembles the act of holding a broom handle and making sweeping motions. In addition, and indeed of more interest, the forms systematically exhibit iconicity for their grammatical class. Relying on the notional definition of noun and verb grammatical class as *thing* and *process*, respectively (Langacker 1987*a*), Wilcox argues that because of their restrained manner and reduplicated movement, ASL noun forms are articulated in a region of space occupied by *things*, and verb forms, because of their salient movement through space, reside in the region of conceptual space occupied by *processes*.

Cognitive iconicity thus defines iconicity not as a relation between the form of a sign and its real world referent but as a distance relation within a multidimensional conceptual space between the phonological and semantic poles of symbolic structures. The typical case for language is that the semantic and the phonological poles of a symbolic structure reside in vastly distant regions of conceptual space. The sound of the spoken word *dog*, for example, has little similarity to the meaning of the word. This great distance in conceptual space between the word's semantic and phonological poles is the basis for Saussure's notion of *l'arbitraire du signe*. When the phonological and semantic poles of a sign lie more closely in conceptual space, arbitrariness is reduced.

It should be noted that metonymy plays a central role in determining these conceptual distance relations. The sound of the spoken word *bow-wow*, for example, does not reside in the same conceptual region as the notion it represents ("dog"); rather, it resides close to the conceptual region occupied by the sound that a dog makes, which is metonymically linked to the concept "dog".

Metaphor also plays a critical role in determining conceptual distances. Metaphor is a mapping across semantic domains (Lakoff and Johnson 1980). Just as a "wormhole" can dramatically alter distances by warping physical space, metaphors are able to change semantic locations in conceptual space. As a result, a semantic location that was formerly distant from a phonological location may be relocated in conceptual space by a metaphorical mapping. If the new semantic location is closer in conceptual space to the sign's phonological location, iconicity is increased.

Cognitive iconicity is able to describe iconic relations in morphology, the lexicon, and the grammar. While the notion of cognitive iconicity is not limited to signed languages, the form of signed languages—objects moving and interacting in three-dimensional space—provides a far richer environment in which cognitive iconicity may be manifest. Much of what we talk about is objects and movements; even when we are not talking of physical objects and movements, we often use metaphor to map abstract notions onto real-world objects and their interactions. For example, it is well known that ideas are often metaphorically conceived as objects. Consciousness may be conceptualized metaphorically as a surface: "He was under general anesthesia for the surgery". In ASL, ideas are often conceptualized as straight objects represented by a straight index finger (P. Wilcox 2000). An ASL word meaning "an idea/thought slipped out of my consciousness" is signed with the extended index finger of one hand situated pointed upright between the index and middle fingers of a B-handshape on the other hand; the index finger is rapidly pulled downward. The visual effect is of an elongated object, represented by the index finger, dropping beneath the surface of consciousness, represented by the flat B-handshape.

29.3.2 Metaphor

Signed language linguists have recently begun to investigate the conceptual process known as metaphor. As we have just seen, signed languages are expressed by handshapes, movements, and location, which themselves carry conceptual import reflected in the overriding iconicity exhibited by signed languages. The strong presence of iconicity in the lexicon of signed languages has created a smoke screen, sometimes leading linguists to confuse metaphor with iconicity.

Brennan (1990: 27), for example, claimed that the British Sign Language (BSL) sign GRASS was an example of a lexical metaphor: "we can see that one set of upright long(-ish), thin(-ish) objects (blades of grass) is represented by another set of upright long(-ish), thin(-ish) objects (fingers)." Rather than exhibiting metaphor, however, this is an example of iconicity. There is no mapping of source domain onto target domain in this sign; the form of the sign GRASS simply resembles its referent, blades of grass. Similarly, Boyes-Braem (1981) originally claimed that the two ASL signs $\mathrm{WRISTWATCH}_1$ (using the F classifier representing the outline of the watch face) and $\mathrm{WRISTWATCH}_2$ (using the L classifier representing the strap circling the wrist) were metaphorical. She now sees them as iconic.

Early research on metaphor in signed language typically found metaphor at the lexical level. Wilbur (1987) was one of the first to systematically explore metaphor in ASL at the lexical level. She noted that many ASL signs exhibit systematic spatialization metaphors. The metaphor HAPPY IS UP is exemplified in signs such as HAPPY, CHEERFUL, and LAUGH. The metaphor NEGATIVE IS DOWN shows up in signs such as LOUSY, IGNORE, and FAIL. Front-to-back spatialization is prominent in the metaphorical mapping of time onto space in ASL: signs such as TOMORROW, NEXT-WEEK, and NEXT-YEAR move forward in the signing space, while YESTERDAY, LAST-WEEK, LAST-YEAR, and RECENTLY move back in signing space.

P. Wilcox (2000) expanded the analysis of lexical metaphor in ASL by demonstrating systematic relationships among the signs used to convey the metaphor IDEAS ARE OBJECTS. For example, within the semantic network created by the superordinate metaphor IDEAS ARE OBJECTS, Wilcox noted several basic level metaphors: IDEAS ARE OBJECTS SUBJECT TO PHYSICAL FORCE, IDEAS ARE OBJECTS TO BE MANIPULATED OR PLACED, IDEAS ARE OBJECTS TO BE GRASPED, and IDEAS ARE OBJECTS TO BE CAREFULLY SELECTED.

Wilcox noted that the metaphors are expressed in ASL by distinct lexical signs. IDEAS ARE OBJECTS TO BE GRASPED and IDEAS ARE OBJECTS SUBJECT TO FORCE may be expressed with a sign beginning with an S-handshape, opening to a 5-handshape. In one narrative, a deaf consultant explained that he had traveled the world collecting the folklore of deaf people. He said that he had committed to memory the hundreds of stories he had seen told by deaf people around the world. His intention was to eventually place them in a historical document. The S-handshape in ASL is used to express the concept of grasping: it would be used to sign “grasp a bicycle handlebar” and is seen in ASL signs meaning “broom” or “to sweep” as well as “motorcycle”. It can also be used to represent grasping a mass of small objects, such as securely holding coins or small stones in one’s hand. In all these cases, the S-handshape represents grasping a physical object. In the narrator’s story, however, the S-handshape is used to represent metaphorical objects—the ideas encompassing the stories. The S-handshape is produced in the location of the signer’s forehead, indicating that the ideas are securely stored in his mind. To convey the concept that the stories will be recorded in a document, the signer simultaneously opens the S-handshape into a 5-handshape representing the release of the ideas, and he moves his hands downward to where a book would be located in front of the signer.

In a similar fashion IDEAS ARE OBJECTS TO BE MANIPULATED OR PLACED uses a handshape commonly used in ASL to represent how one would grasp a flat object such as a piece of paper or a book. This handshape appears in a set of signs used to express the manipulation of objects: moving an object from one location to another (for example, “to give”), removing a book from a shelf, decorating a Christmas tree, and so forth. When ideas are metaphorically understood to be objects to be manipulated, moved, or placed, this handshape is used. For example, a signer would use this handshape to talk about moving ideas around various locations in her mind to convey the concept of organizing her thoughts.

When an idea is metaphorically discriminated, carefully selected, or extracted from some location, the F-handshape is used. This handshape is used in ASL to convey the notion of selecting or picking up small physical objects such as seeds, small buttons, or a sewing needle. When used metaphorically with the concept of ideas, it suggests that the idea is being carefully selected. It also implies limited quantity: whereas the S-handshape may represent the grasping of many ideas, when the F-handshape is used metaphorically in connection with ideas, it suggests that only a single idea has been selected.

Front-back spatialization can be applied creatively to the metaphor IDEAS ARE OBJECTS. In ASL, the front of the head is used to convey the concept of conscious thinking, as exemplified by signs such as REMEMBER, THINK, MULL-OVER, WONDER, IMAGINE, KNOW, UNDERSTAND, MEMORIZE, INVENT, and WISE (P. Wilcox 2000). A signer can use front-back spatialization to create a novel sign, for example by articulating REMEMBER at the back of the signer's head, meaning "unconsciously remember".

While the congruent metaphors MIND IS A CONTAINER and IDEAS ARE OBJECTS (that are contained within the mind) are pervasive in signed languages, including ASL, British Sign Language (BSL), Catalan Sign Language (LSC), French Sign Language (LSF), and Italian Sign Language (LIS), they are not universal. Cultural variability in metaphors used among the world's signed languages has received only cursory attention. In Japanese Sign language (JSL) the metaphor TORSO IS A CONTAINER interacts with IDEAS ARE OBJECTS and UNDERSTANDING IS CONSUMING FOOD. In JSL, the torso is understood as the container for ideas; the body, the same cavity where food is digested, instead of the forehead, provides a container where understanding takes place. The JSL sign meaning "to comprehend" is DRINK-QUICK. British Sign Language uses a unique metaphor, UNDERSTANDING IS DRAWING THE STRING OF A BOW, which may reflect a cultural experience.

29.3.3 Metonymy

From the earliest contemporary studies of signed languages, linguists have noted the presence of metonymy (Stokoe et al. 1965; Mandel 1977; Wilbur 1987). As was the case for metaphor, these early studies of metonymy predominantly focused on simple lexical metonymy. Wilcox et al. (2003) present a typology of lexical metonymy with data from ASL and LSC. Types of metonymies reported include:

- **Prototypical characteristic for whole entity**. In both ASL and LSC the signs for "bird", "horse", and "cow" depict prototypical physical properties of these animals: the beak, the ears, and horns, respectively.
- **Action for instrument**. Signs in ASL and LSC meaning TYPEWRITER, GUITAR, TOOTHBRUSH, and OAR exemplify this metonymy. In the ASL sign TYPE-WRITER, for example, the hands and fingers are moved in a way representing the action of typing.
- **Prototypical action for activity**. The hands and their movement may be used to represent some prototypical action taken with an object; this in turn may come to metonymically express the general activity. In LSC, the signs DRINK-BEER, DRINK-BRANDY, DRINK-RUM-AND-COKE use specific handshapes

representing interaction with a container of a specific, prototypical shape, as well as movements characteristic of drinking from these containers.

- **Salient characteristic of a specific person for general quality**. A number of signs in LSC rely on a type of iconic, gestural metonymy in which a salient characteristic of a well-known person is extended to stand for a more general quality. These metonymies also typically involve metonymic chains. For example, the LSC sign CHARLIE-CHAPLIN is a compound that iconically depicts Chaplin's moustache and the movement of holding a cane and moving it in circles as Chaplin did. Similar examples from LSC include HITLER (the sign iconically depicts Hitler's characteristic moustache) for "bad" or "evil"; DALI (depicting Dali's characteristic moustache) for "crazy"; and JESUS-CHRIST for "suffering person".
- **Deviant behavioral effect for intensity of experience**. The LSC sign that we gloss as CRAZY-EYES (an iconic sign depicting the eyes open wide and moving in wild circles) means "really good"; the sign could be used, for example, to describe delicious food. Similarly, OPEN-MOUTH means "astonishment"; and APOPLEXY (iconically depicting the wild movements of a person experiencing a seizure) can be used to describe any "incredible" experience.

Signs often exhibit complex metonymic chains. In Saudi Sign Language, there is a sign made with a handshape that represents the handle of the Arabic curved scimitar or saif (prototypical characteristic for whole entity). The sign's movement represents putting the saif into a sash or belt. This complex metonymic sign denotes the city of Bahrain for the men in that city who would typically wear saifs.

The ASL sign THINK-HEARING demonstrates the complex way in which metonymy, iconicity, and metaphor interact (P. Wilcox 2000). Etymologically, THINK-HEARING derives from the sign SAY, which is articulated at the mouth with tiny circular movements that iconically indicate the flow of speech from the person who is talking. SAY is also metonymic because the circling movements iconically represent the breath emanating from the speaker's mouth. The exhaled air is metonymically extended to stand for the speech produced by the person.

In a semantically extended sense, SAY has also come to denote a hearing person, one who speaks. The circling movements that represent speech are an example of synecdoche, where a part (the act of speaking) stands for the whole (the hearing person doing the speaking). This metonymy is then extended when the sign HEARING-PERSON is used also to represent the thoughts and culture of hearing people.

When HEARING-PERSON is placed at a different location, multiple metonyms are further formed by what Goossens (1990: 338) calls cumulative metaphtonymy—a metaphor derived from metonymy. In this case, the sign is moved from the mouth to the forehead. By virtue of this simple change in location, HEARING-PERSON becomes THINK-HEARING. The metonymic expression THINK-HEARING takes on a metaphoric mapping. The sign no longer metonymically refers to voice production, to a hearing person, or even to the extended metonym for the culture and values of hearing people. The sign SAY, through these metonymic and metaphorical extensions as THINK-HEARING, comes to denote a deaf person, but one who thinks like a hearing person, who rejects signing ASL, accepts speech and speech-related signing, and uncritically adopts the ideology and cultural values of the hearing world.

We have seen that signed languages differ in their basic metaphors. An interesting metaphorical category found in LSC is IDEAS ARE LIQUID. When an idea or thought is learned from other people, the lexical expression used in LSC reproduces a movement one would make when sipping liquid through a straw. The agent sips in the learned material, moving the two-handed straw-like classifier in an arc from left to right, sucking in the "ideas and thoughts" of other people. This metaphor interacts with metonymy in LSC. The straw itself is not directly represented; rather, the fingers that hold the straw are iconically and metonymically (simple synecdoche) depicted.

Metonymy has been implicated in diachronic change in many languages. Sweetser (1990: 35), for example, reports: "In all Indo-European languages, the verb meaning 'feel' in the sense of tactile sensation is the same as the verb indicating general sensory perception." Wilcox and Wilcox (2003) have reported that in ASL, the sign FEEL passes from the physical into the non-physical realm. The first extension leads from touching to physical sensing or feeling: the sign TOUCH is phonologically related, using the same handshape as FEEL. A second extension leads from physical sensing to internal reaction (feeling emotion). Finally, we also see an extension from emotion to cognitive or mental reaction (belief, judgment, and decision).

Although the cognitive processes of iconicity, metaphor, and metonymy have received decades of study in signed languages, linguists are still finding unique forms of expression made possible by the visual-gestural modality of these languages. Clearly, much more research on these cognitive processes and their complex interactions in conventional and novel or creative expression in signed languages is needed.

29.4 Grammaticalization

Grammaticalization operates in signed languages much the same as it does in spoken languages, with some intriguing differences. One example is the grammaticalization of modal forms. Just as linguists have found that deontic and epistemic modals typically develop out of lexical forms, researchers have discovered that ASL modal auxiliaries often develop from lexical forms; for example, the ASL modal CAN developed historically from an ASL lexical form meaning "possessing physical strength" (Wilcox and Wilcox 1995).

Another example where we see similar grammaticalization paths across spoken and signed languages is the development of a completive or perfective marker from a lexical source meaning "finish". In addition to occurring in language groups as diverse as Bantu, Niger-Congo, Andean-Equatorial, and Tibeto-Burman (Bybee et al. 1994; Heine and Reh 1984), the same development has been described for the sign FINISH in ASL (Janzen 1995).

Unique differences in grammaticalization of signed languages are found in two areas. First, grammaticalization may be extended in the case of signed languages to account for the development of lexical and grammatical material, both manual and facial, from gestural sources. It has been proposed that ASL deontic and epistemic modal forms developed from gestural sources (Wilcox and Wilcox 1995); that the ASL future marker developed from a lexical form meaning "to leave" which had its source in a pan-Mediterranean gesture signaling departure (Janzen and Shaffer 2002); and that the ASL negative ability marker "cannot" can be traced to the deontic modal "must" and ultimately to a gesture indicating financial obligation (Shaffer 2002). The second difference is that in certain cases, grammatical forms in signed languages develop out of gestural forms without passing through a lexical stage. In these instances, the gestural forms enter the linguistic system first as prosody or intonation, and then become grammatical forms. For example, the gestural source for topic marking is likely a questioning facial gesture of eyes wide open. The communicative facial expression came to mark polar questions (see Zeshan 2004), which then further grammaticalized to mark topics (Janzen 1999).

This finding is strikingly different from the evidence from spoken languages. As Heine and Kuteva (2007: 108) point out, in spoken languages we invariably find that the first stages of grammatical evolution are restricted to lexical categories. They note that, in the case of signed languages, gesture may provide an alternative pathway that bypasses the lexicon, "leading straight from manual or non-manual gesture to functional marker" (ibid.: 109). Heine and Kuteva propose that more detailed analysis of suprasegmental grammatical forms such as intonational contours of polar question marking could strengthen this case. We hope to provide some of this detail, at least for the second route of grammaticalization, in section 29.5.3.

29.5 Gesture and signed languages

The relation between gesture and signed languages has recently emerged as a significant topic in the analysis of signed languages. For centuries, signed languages were regarded as nothing more than gesture. The nadir of this view was reached in the late 1800s when deaf educators who wished to forbid the use of signed language in deaf education declared that "Gesture [i.e., signed language] is not the true language of man which suits the dignity of his nature. . . . Thus, for us it is an absolute necessity to prohibit that language and to replace it with living speech, the only instrument of human thought" (Giulio Tarra, quoted in Lane 1984: 393–4).

Because signed languages were commonly equated with gesture, the relation between gesture and signs was for decades a taboo research topic among linguists motivated to demonstrate the linguistic status of signed languages. The first break came with the publication of *Gesture and the Nature of Language* (Armstrong et al. 1995), which claimed that gesture is implicated in the evolutionary development of language. While continuing to explore the evolutionary link between gesture and language (Armstrong and Wilcox 2007), Wilcox has also gone on to study the diachronic process by which gesture enters the linguistic system.

29.5.1 The gesture–signed language interface

Wilcox (2004*b*; 2005; 2007) has argued that gestures follow two routes as they codify and grammaticalize. The first route begins with a gesture that is not a conventional unit in the linguistic system. This gesture becomes incorporated into a signed language as a lexical item. Over time, these lexical signs acquire grammatical function (Figure 29.1).

The second route proceeds along a different path. The source in this route is one of several types including the manner of movement of a manual gesture or sign, and various facial, mouth, and eye gestures. The second route follows a path of development from gesture to prosody/intonation to grammatical marker (Figure 29.2). Notably, the second route bypasses any lexical stage.

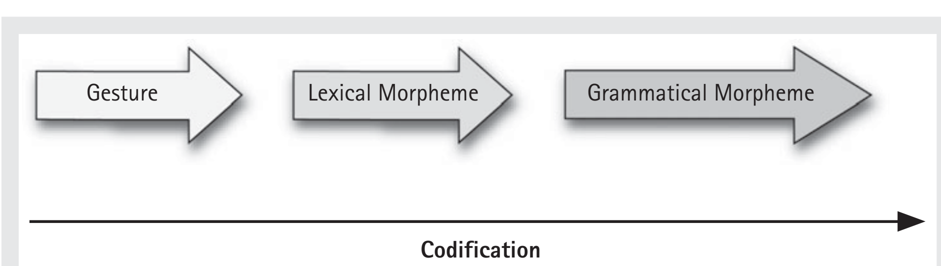


Figure 29.1. First route from gesture to language

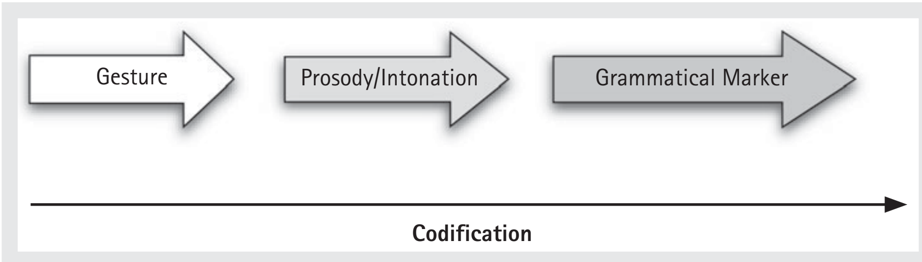


Figure 29.2. Second route from gesture to language

29.5.2 The first route: From gesture to word to grammatical morphology

Four sources of evidence for the developmental path leading from gesture to lexical morpheme to grammatical morpheme are presented here: futures, venitives, markers of obligation, and evidentials and epistemic modals.

Data from a cross-section of the world's spoken languages demonstrate that there are three common sources for future markers: desire, obligation, and movement verb constructions (Bybee et al. 1994). Lexical morphemes meaning "come", "go", and "desire" are a common source of grammatical morphemes used to indicate the future in many spoken languages. Using a corpus of historical as well as modern conversational data, Shaffer (2000) and Janzen and Shaffer (2002) have demonstrated that the grammatical form used to mark future in ASL (Figure 29.3a) developed from the lexical morpheme "go" (Figure 29.3b).

The gestural source of the future marker is described by de Jorio (2000 [1832]). It is produced with the palm of the hand open and held edgewise, moved upward several times. Morris et al. (1979) identify this as a gesture still in use among hearing people in the Mediterranean region to signal departure-demand and departure-

(a) (b)

Figure 29.3. FUTURE and GO in American Sign Language

Figure 29.4. French gesture meaning "departure"

description (Figure 29.4, from Wylie and Stafford (1977)). The gesture appears in LSF as the lexical morpheme PARTIR "depart" (Figure 29.5, after Brouland (1855)).

Another set of examples documenting the first route comes from venitives, gestures signaling movement toward speaker. This path begins with a gesture meaning roughly "come here" identified by de Jorio as CHIAMARE, "to call or summon someone": "Fingers extended and then brought towards the palm several times" (de Jorio 2000 [1832]: 124).

The "come here" gesture appears as a lexical item in a number of signed languages, especially those used in the Mediterranean region. This form appears in ASL in a variety of senses including requests for physical movement, incitement to action, and requests for metaphorical movement such as the transfer of information or ideas. A signer might use an ASL lexical sign derived from the "come here" gesture to request that more information be provided. For example, when a deaf consultant was asked how she became interested in linguistics, she replied, "I took a beginning course and became fascinated with linguistics—'Give me more!"' where the phrase translated here as "Give me more!" was the two-handed ASL lexical sign COME-HERE. Higgins (1923) gives the form as NECESSITY (Figure 29.6).[2]

[2] This use is no longer attested among ASL users.

Figure 29.5. PARTIR in Old French Sign Language

In Catalan Sign Language (LSC) the "come here" form appears as a lexical sign to request physical movement or, more generally, an invitation to join or affiliate with a group. It also appears in a more specific sense as the lexical sign EMERGÈNCIA "emergency". In Italian Sign Language (LIS) the form functions to request physical

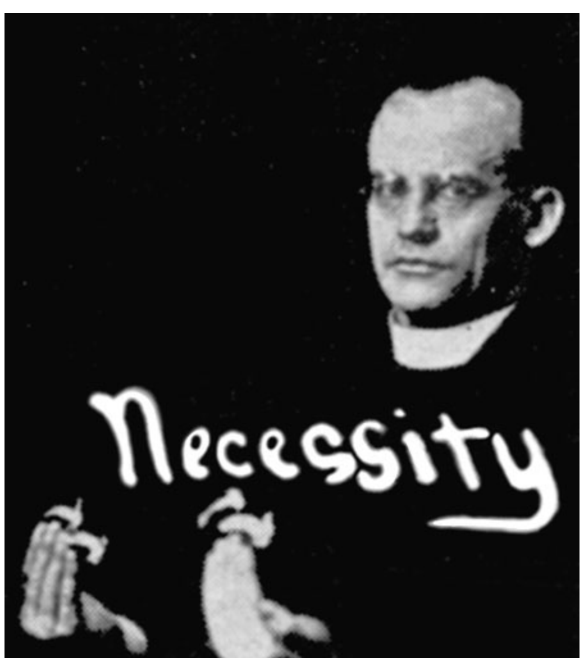


Figure 29.6. NECESSITY in American Sign Language (archaic)

movement. The “come here” form also is used in LIS to encourage action on the part of the interlocutor. For example, in one recorded LIS conversation, a deaf teacher was asked whether hearing students learning LIS could be forced to sign. She responded that students should be encouraged rather than forced to sign in class, using the LIS one-handed COME-HERE form to mean “encourage”.

The meanings of the gestural and linguistic forms are metonymically related and motivated by pragmatic inferences (Traugott and König 1991; Panther and Thornburg 2003) and metaphor (Heine et al. 1991). Pragmatic inferencing motivates the extension from a request for physical movement to necessity and emergency: one reason I might request that another person come to me is because I need them. The extension from a request for physical movement to a request for information is metaphorically motivated by mapping the movement of physical objects toward the speaker onto metaphorical objects of communication (Reddy 1979). An inferential link motivates the extension to encouragement: one reason I might request you to produce a behavior (metaphorically, “send” an object from you to me) is because I want to encourage you.

The third set of data documenting the first route comes from the development of obligation verbs. Shaffer (2002) notes that the ASL deontic modal MUST (the forefinger is bent into a “hook” shape, and the hand is oriented so the palm faces down; push the hand downward by bending it at the wrist) is related to the French Sign Language form IL FAUT “it is necessary” (the forefinger is straight, and the hand is oriented to the ipsilateral side; push the hand downward by twisting the forearm). IL FAUT is also attested in mid-nineteenth-century LSF (the extended index finger is directed down toward the ground). It is likely that these forms derive from a gesture used as early as Roman times to signal obligation. Dodwell (2000: 36) discusses a gesture that he calls an imperative: “It consists of directing the extended index finger towards the ground.” According to Dodwell, the gesture was described by Quintilian in the first century AD: “when directed towards the ground, this finger insists” (ibid.).

Wilcox and Wilcox (1995) identified epistemic and evidential modal forms in ASL that developed from lexical morphemes having gestures as their source. The ASL evidential forms SEEM, FEEL, and CLEAR/OBVIOUS grammaticalized from lexical morphemes MIRROR, FEEL (in the physical sense), and BRIGHT, respectively. Each of these lexical morphemes can be traced in turn to a gestural source. The full developmental path for these forms is:

1. [gesture enacting looking in a mirror] > MIRROR > SEEM (evidential)
2. [gesture enacting physically sensing with finger] > FEEL (physical) > FEEL (evidential)
3. [metaphorical gesture indicating rays of light] > BRIGHT > CLEAR/OBVIOUS (evidential)

In each case the path is from gesture to lexical morpheme to grammatical (modal or evidential) morpheme.

Data from LSC (Wilcox et al. 2000) also demonstrates the emergence of grammaticalized modal and evidential forms from gestural and lexical sources. The LSC forms EVIDENT, CLAR, PRESENTIR, and SEMBLAR (Figure 29.7a–d) have developed subjective senses which encode the agent's expression of himself or herself in the act of utterance. This tendency for meanings to become based in speaker subjectivity is one indication that a form has become more grammatical (Traugott 1989).

As a lexical morpheme EVIDENT has a range of physical senses denoting visual perception, including intensity of color; prominent or salient, such as a person who stands out because of her height; "sharp, well-defined", such as indicating sharpness of an image; and "obvious", as when looking for an object located in front of you. As a grammatical morpheme EVIDENT denotes subjective, evidential meanings such as "without a doubt", "obviously", and "logically implied". The lexical morpheme CLAR is used in more concrete meanings to denote "bright" or "light". It may also be used in a more abstract sense to denote clear content, a person's skill in signing, or ability to explain clearly. As a grammatical morpheme CLAR encodes speaker subjectivity and may be used in the same context as EVIDENT.

Used as a lexical morpheme, PRESENTIR denotes the sense of smell. The grammatical morpheme PRESENTIR is used to express the speaker's inferences about actions or intentions. For example, an LSC signer might state her inference that, while someone has said he would go on a trip, she suspects that he will not. In this context, she expresses her suspicion with PRESENTIR, literally "it smells" like the person will not go. When used as a lexical morpheme SEMBLAR denotes physical resemblance. The grammatical sense of SEMBLAR may be used to express the speaker's subjective belief that an event is or is not likely to occur.

As we saw for the ASL data, these LSC forms have sources in metaphorical or enacting gestures indicating the eyes and visual perception (EVIDENT), bright light (CLAR), the nose and the sense of smell (PRESENTIR), and physical, facial appearance (SEMBLAR). Once again, the full developmental path is from gesture to lexical morpheme to grammatical morpheme.

29.5.3 The second route: From gesture to prosody/intonation to grammatical marker

The second route follows a distinctly different pathway from the first route. It begins with either the manner of movement of a manual gesture, or certain facial gestures. These nonlinguistic gestures enter the linguistic system not as lexical forms but as prosody and intonation; as they grammaticalize, they take on grammatical function.

(a) EVIDENT

(b) CLAR

(c) PRESENTIR

(d) SEMBLAR

Figure 29.7. Evidential forms in Catalan Sign Language

Prior research by Dwight Bolinger (1986) suggests a link between gesture and prosody-intonation. Bolinger held that intonation is part of a gestural complex whose primitive and still surviving function is the signaling of emotion. He also suggested that gesture and intonation develop out of expressive origins and then take on more codified linguistic behavior.

29.5.3.1 *Prosody and intonation in signed languages*

Friedman (1977) was one of the first researchers to document the expression of prosody and intonation in signed languages. She observed that signs marked with emphatic stress are larger, tenser, faster, and with longer duration than unstressed signs. Other differences in stressed vs. unstressed signs include changes in the manner of production, both in rhythmic characteristics (addition of tension, restraint, or faster movement), and in the movement itself.

Wilbur and Schick (1987) observed that in spoken languages, the primary cues for linguistic stress are increased duration, increased intensity, and changing the fundamental frequency. Markers of increased duration include larger movement, slower movement, repetition, and added movement. Markers of increased intensity include the addition of non-manuals (for example, eye or mouth gestures), sharp boundaries between signs, higher articulation of signs in the signing space, increased tension of articulation, and more forceful articulation.

Examining data from Israeli Sign Language, Nespor and Sandler (1999) found that the phonetic correlates of prosodic prominence included reduplication, a hold at the end of the prominent sign, and a pause after the last word of the phonological phrase. Kingston (1999) predicted that signs may be made phrasally prominent by increasing the size, speed, and/or acceleration of their movement. Once again, we find that the phonetic correlates to prosody in signed languages predominantly lie in manner of movement: reduplication, hold (stopping the movement), speed, and acceleration. Increased size is also related to movement, since decreasing/increasing the size of a sign typically results in faster/slower movement, respectively. Sandler (1999) maintains a distinction between prosody and intonation in signed languages, suggesting that facial articulations may be best understood as fulfilling the role of intonation, and proposes that the primitives of intonation are different positions of the brows, eyes, cheeks, mouth, and head.

In summary, the phonetic correlates of prosody and intonation in signed languages appear to be (1) *facial articulations*, and (2) changes to a sign's movement (speed, acceleration, duration, repetition, size, tension, force), which we group under the category *manner of movement*.

29.5.3.2 *Facial articulations*

Two striking features characterize nonmanual or facial articulators. First, facial articulations are predominantly used to code grammatical functions, such as topic, interrogatives, and imperatives. Facial articulators rarely if ever are the sole means of expressing lexical morphemes. Second, these markers serve remarkably similar function across a wide range of genetically and geographically unrelated languages. In a typological study of interrogatives, Zeshan (2004) found that all 30 languages studied used non-manual marking for polar questions. In addition, Zeshan reported that the non-manual signals used to mark polar questions tend to be quite similar across signed languages, typically involving a cluster of facial gestures including eyebrow raise, eyes wide open, eye contact with the addressee, head forward position, and forward body posture.

29.5.3.3 *Verb aspect as grammaticalized manner of movement*

Klima and Bellugi (1979) described alternations in the manner of movement used to mark verb aspect in ASL. Klima and Bellugi called these alternations inflectional morphology, implying that they are highly codified grammatical forms. This would suggest that the grammaticalization of prosody results in grammatical markers expressed by manner of movement (Wilcox 2005). Evidence for the grammaticalization of prosody as verb aspect also comes from LIS. Pizzuto (2004 [1987]) observed that temporal aspect can be expressed in LIS both lexically and by means of systematic alterations of the verb's movement pattern. Manner of movement of the verb form can indicate "suddenness" of an action by means of a tense, fast, short movement (e.g., the distinction between "to meet" and "to suddenly/unexpectedly meet" someone); a verb produced using an elongated, elliptical, large and slow movement specifies that an action is "repeated over and over in time" or "takes place repeatedly in time".

29.5.3.4 *From gesture to prosody/intonation to grammatical function*

We suggest that these data document developmental points along the second route. Manner of movement and facial articulations begin as gesture, but of a type that enters the linguistic system as prosody and intonation, bypassing the lexical stage. In early stages of codification they continue to exhibit a high level of gradience and often serve to mark speaker attitude.

As they further codify, manner of movement and facial articulations take on more grammatical functions, finally appearing as question, topic, and conditional markers; adverbial forms; and verb aspect. As described in section 29.3.1, manner of movement is also used to mark noun–verb distinctions in ASL (Supalla and Newport 1978). The grammaticalization of manner of movement also marks strong vs. weak modal forms (Wilcox 2004*b*), as well as deontic vs. epistemic modality (Wilcox and Shaffer 2006).

The gestural heritage of these forms provides an explanation for Zeshan's (2004) finding that manner of movement and facial articulations perform similar functions across genetically and geographically unrelated languages. Although it is certainly true that emotion is not signaled in a uniform way across cultures, it is nevertheless the case that the manual and facial gestural inventory for marking emotion is much more restricted and similar than the crosslinguistic inventory of emotion words. Cultural differences notwithstanding, we all recognize a gesture made in anger, or a face that signals sadness. It is not surprising that, when gestures of this type make their way into signed languages as prosody and intonation, they are used to express the same sorts of general semantic notions and follow roughly similar grammaticalization paths crosslinguistically.

29.6 Summary

Compared to spoken language linguistics, the analysis of signed languages is still in its infancy. Even so, linguists have contributed a great deal to our knowledge of the history, grammars, and typology of these natural languages. The study of signed languages provides insights into many of the longstanding questions pondered by linguists, such as the nature of duality of patterning and the origin and evolution of the human language faculty. The analysis of signed languages permits linguists and cognitive scientists to address important questions about the nature of grammaticalization, metaphor, iconicity, and metonymy. Signed languages also open up new horizons for linguistic analysis. Because signed languages are visually perceived, they extend the linguist's scope of study by adding the capability of triangulating optical with acoustic data and observing first-hand how material that starts its life outside of the linguistic system becomes conventionalized, lexicalized, and grammaticalized.

CHAPTER 30

SIMPLER SYNTAX

PETER W. CULICOVER

A syntactic theory for natural language provides an account of the correspondences that hold between phonology and meaning. The core idea of Simpler Syntax (henceforth SS) is expressed by the Simpler Syntax Hypothesis.

Simpler Syntax Hypothesis (SSH): The most explanatory theory is one that imputes the minimum syntactic structure necessary to mediate between phonology and meaning.

The SSH is based on what is generally accepted as good scientific practice: if we have an indispensable mechanism at our disposal, we should determine whether or not it is capable of accounting for the observed phenomena before we introduce new mechanisms. Culicover and Jackendoff (2005) argue that semantic representation in particular offers a rich source of explanation that makes possible a dramatic simplification of syntactic representations and of syntactic theory.

The SS approach contrasts with that of mainstream generative grammar (MGG), which has been driven in large part by the goal of formulating maximally uniform syntactic accounts of natural language phenomena. SS argues that the richness of the syntax/semantics correspondence makes syntactic uniformity unnecessary, and that in many cases a strict application of uniformity yields incorrect or incomplete analyses of a number of important phenomena.

This chapter is the product of joint work with Ray Jackendoff. I take full responsibility for any errors. I am grateful to Anne Abeillé and the Department of Linguistics of the Université de Paris VII for providing me with the opportunity to lecture on this material during February and March 2008.

In this chapter I summarize briefly the architecture of SS, and review some of the implications of the SS approach for the analysis of phenomena such as ellipsis, control and raising, argument alternations such as active/passive, and A′ constructions. It will be seen that much of SS is not new. SS should not be seen as a radical theoretical departure but rather as a return to an earlier and, we believe, correct perspective about how language works. The overlaps between SS and the theory of syntax outlined in Chomsky's *Aspects* (Chomsky 1965) as well as Head-Driven Phrase Structure Grammar (HPSG) and Lexical-Functional Grammar (LFG) are substantial and by no means accidental.

30.1 An overview of Simpler Syntax

30.1.1 Architecture

SS assumes that a grammar consists of the following:

a. structural principles governing phrase structure, morphology, and agreement;
b. conceptual structure;
c. phonological structure;
d. rules that state correspondences between the three representations a–c;
e. the primitive grammatical functions (GFs) Subject and Object that mediate correspondences between parts of conceptual structure and syntactic constituents;
f. a lexicon that catalogues the unpredictable interfaces between phonological, syntactic, and semantic structures for particular words and phrases;
g. functional principles that govern the alignment of syntactic structure with prosody and information structure.

The correspondence rules (d) and (a)–(c) implement the Parallel Architecture of Jackendoff (1997); Jackendoff (2002)—see Jackendoff (this volume) for the details. The GFs are assumed to be primitives of the language faculty; they are realized through a range of configurational and morphological devices in natural languages. They constitute a tier that regulates the mapping between CS arguments and NP arguments in syntax. This tier is designed along the lines of a subset of LFG, and has counterparts in Relational Grammar (RG), HPSG (the complement hierarchy), and mainstream generative grammar (abstract Case). The lexicon incorporates not only the correspondences for individual words of a language (that is, their phonological, syntactic, and semantic properties), but the correspondences for more complex expressions, such as idioms and constructions—see Jackendoff's chapter in this volume for some further details. I will have nothing to say here about (f).

30.1.2 An antidote to uniformity

The most significant principle of SS is that it abandons syntactocentrism (the generation of all linguistic combinatoriality through the syntactic component) and the pervasive syntactic uniformity methodology of MGG. The consequences are far-reaching. I give a few examples to show what happens when we give up syntactic uniformity. There are three types of uniformity applied in MGG (although they often have overlapping consequences).

Structural Uniformity (SU): An apparently defective or misordered structure is actually a distorted regular form. For example, it follows from SU that the direction of branching is uniform in natural language, and that variations in constituent order are the consequence of movements (Kayne 1994).

Interface Uniformity (IU): The syntax–semantics interface is maximally simple, in that meaning maps transparently into syntactic structure; and it is maximally uniform, so that the same meaning always maps into the same syntactic structure. It follows from IU that the assignment of thematic roles is uniform (i.e., UTAH (Uniformity of Theta Assignment Hypothesis), Baker 1988: 46), and from this it follows that the syntactic structure of a controlled predicate is that of a full sentence with an invisible (i.e., phonologically null) subject. Similarly, it follows that sentences with ellipsis will have an analysis in which they are assigned an invisible syntactic structure that is rich enough to account for their full interpretation.

Derivational Uniformity (DU): Where possible, the derivations of sentences are maximally uniform. It follows from DU that sentences with gaps that appear to form part of an A′-chain (e.g., wh-questions, topicalization, relative clauses, *tough* constructions, *too/enough*) will all have derivations in which the gap is produced in the same way, through A′ movement (Chomsky 1977).

SS proposes that for each characteristic of natural language that uniformity is intended to account for, a non-uniform account is syntactically simpler, and at least descriptively equivalent, if not more adequate. A crucial part of the exercise is to show that MGG accounts are often initially appealing because they fail to take into account the full range of phenomena. For example, in the case of IU, it is possible to account for a subset of the cases of ellipsis in terms of invisible syntactic structure that is identical to the overt syntactic structure of some antecedent. But more complex cases require that semantic and pragmatic devices be employed in the interpretation of the elliptical construction. Given the availability and necessity of such devices, the case for the syntactic account is substantially weakened. Some examples are given below to illustrate.

30.1.3 Consequences for grammatical theory

I summarize here the main consequences of abandoning uniformity. Without SU, word order variation must be represented directly in the grammar. This means that linear order must be a primitive of the syntactic theory and not a derived property. (This is not problematic, given the epistemological priority of linear order.) Without SU, phrase structure need not be as uniform as is required by a strict version of X′ theory, and in fact exocentricity is in principle possible. Explicit statements about word order allow it to be accounted for without assuming invisible functional heads, uniform spec-head agreement, and the discharge of (strong) features.

Abandonment of IU means abandonment of UTAH. Without UTAH, phrases that have a full propositional interpretation, such as infinitivals, gerunds, and secondary predicates, get their interpretation through rules of interpretation. Moreover, argument alternations, such as active/passive, raising to subject, raising to object, active/antipassive, applicatives and causatives, are not derived from uniform syntactic structures but are alternate realizations of semantic argument structures. Hence there is no PRO, no syntactic control theory, and no A movement. These are dispensed with in SS. The cost is that there must be rules of interpretation that allow for alternative mappings of argument structure to the syntax. But such rules are independently necessary to account for non-systematic lexically governed alternations.

Moreover, without IU, there is no motivation for A′ movement. A constituent in an A′ position does not have to "move" from an underlying canonical position in order to acquire its interpretation—it can do so through its relationship to the gap corresponding to the canonical position. There are ways to generate gaps without movement. Hence in SS there is no A′ movement. The cost is that there must be rules of interpretation for A′ chains. But such rules are independently necessary for chains that do not involve gaps.

Abandonment of DU means that there are no movements "at Logical Form (LF)", or movements of invisible bundles of features. The cost is that the scope of wh in situ and quantifiers is supplied by rules of interpretation and not read off of abstract syntactic representations. A further consequence is that there is no possibility of invisible head-to-head movements "at LF". Sections 30.2, 30.3, and 30.4 develop these consequences for some representative cases.

30.2 Simpler Syntax on ellipsis

There are many kinds of ellipsis. I focus here on bare argument ellipsis (BAE) and its interrogative counterpart, sluicing. BAE is exemplified in B's reply to A in example (1).

(1) A: Harriet's been drinking.
B: Yeah, scotch.

B's reply conveys the same meaning as sentence (2), thus going beyond the meanings of *Yeah* and *scotch*.

(2) B: Yeah, Harriet's been drinking scotch.

The question that must be resolved is: What is responsible for the interpretations and the forms? There are two basic alternatives:

(i) the fragment has a full syntactic structure. The full syntactic structure accounts for the observed form and the meaning of the fragment.
(ii) it does not have a full syntactic structure. The fragmentary syntactic structure and the interpretation are licensed through the relationship between the fragment and the antecedent.

Evidence in support of (i) consists of the fact that the BA (bare argument) behaves syntactically and morphologically as though it is embedded in a complete syntactic structure. The following are classic examples from German that illustrate this. The verb *folgen* "follow" assigns dative case to its direct object, while *suchen* "seek" assigns accusative case to its direct object. The dative case is required on *wem* "who" and *dem Lehrer* "the teacher" in (3a), while the accusative case is required on *wen* and *den Lehrer* in (3b).

(3) a. A: Wem/*Wen folgt Hans?
who-DAT/who-ACC follows Hans
'Who is Hans following?'

B: Dem/*Den Lehrer.
the-DAT/the-ACC teacher
'The teacher.'

b. A: Wen/*Wem sucht Hans?
who-ACC/who-DAT seeks Hans?
'Who is Hans looking for?'

B: Den/*Dem Lehrer.
the-ACC/the-DAT teacher
'The teacher.'

The cases are accounted for if it is assumed that (3a.B) contains an invisible *folgt*, while (3b.B) contains an invisible *sucht*.

A similar situation can be found in English preposition selection in oblique arguments. The following examples show that *flirt* selects *[$_{PP}$ with NP]*, *proud* selects *[$_{PP}$ of NP]*, nominal *pride* selects *[$_{PP}$ in NP]*, and verbal *pride (oneself)* selects *[$_{PP}$ on NP]*.

(4) a. A: I hear Harriet has been flirting again.
B: i. Yeah, with Ozzie.
ii. *Yeah, Ozzie.

b. A: John is very proud.
B: Yeah, of/*in/*on his stamp collection.
[cf. proud of/*in/*on NP]

c. A: John has a lot of pride.
B: Yeah, in/*of/*on his stamp collection.
[cf. pride in/*of/*on NP]

d. A: ?John really prides himself a lot.
B: Yeah, *in/*of/on his stamp collection.
[cf. pride oneself *in/*of/on NP]

The preposition selection can be accounted for if the appropriate lexical heads are assumed to be present, but invisible, in the syntactic structure that contains the PPs in B's responses.

Sluicing is exemplified by examples such as the following. The key characteristic of sluicing is that only the wh-phrase appears, yet the sentence has the interpretation (in context) of a wh-question.

(5) Harriet said something nasty, but I don't remember what$_i$ ['she said t_i']

MGG approaches to sluicing (e.g., Merchant 2001) assume that a constituent is deleted under identity with a constituent in the antecedent; in this case, the deleted constituent is *she said* t_i. The strongest argument for a syntactic account of sluicing is that there is a variant[1] in which a preposition can be stranded. (6a) is clearly understood as meaning the same as (6b).

(6) a. John went to NY with someone, but I couldn't find out who with.
b. John went to NY with someone, but I couldn't find out who$_i$ John went to NY with t_i.

As Merchant points out, languages that lack preposition stranding allow sluicing but do not allow this stranding variant. English, which allows preposition stranding, has this construction.

In spite of this prima facie evidence for a syntactic account of BAE and sluicing, there are difficulties. Most significantly, both for BAE and sluicing, it is assumed in MGG that the overt fragment is moved from its canonical position, and the resulting constituent is deleted under identity. But many well-formed cases of BAE and sluicing do not correspond to well-formed sentences when the supposed deleted constituent is overt. The following examples, based on Culicover and Jackendoff 2005, illustrate the problem for BAE. They violate constraints banning

[1] Called "slifting" by Ross 1973 and "sluice-stranding" by Culicover 1999.

extraction from islands (7), and from proper names (8), compounds (9)–(10) and words (11).

(7) a. A: What kind of scotch does Harriet drink?
B: Expensive. [= 'Harriet drinks expensive scotch']
b. *Expensive$_i$, Harriet drinks [t_i scotch]. [Left branch condition]

(8) a. A: Did Susan say that she saw PAT Smith?
B: No, KIM.
b. *Kim$_i$, Susan said that she saw [t_i Smith].

(9) a. A: Let's get a pizza.
B: Pepperoni?
b. *Pepperoni$_i$ let's get [a t$_i$ pizza].

(10) a. A: Is that a Navy flight suit?
B: No, Army.
b. *Army$_i$, that is a [t_i flight suit].]

(11) a. A: Is Sviatoslav pro-communist or anti-communist these days?
B: Pro.
b. *Pro$_i$, Sviatoslav is [t_i-communist] these days.

Ross 1969 made similar observations regarding sluicing. For example, (12a) is well-formed, but (12b) is a violation of the complex NP constraint.

(12) a. This is a picture that shows Harriet drinking something, but I'm not sure what$_i$.
b. *This is a picture of Harriet drinking something, but I'm not sure what$_i$ this is a picture that shows Harriet drinking t_i.

It is a simple matter to stipulate that constraints on extraction are not in effect when the gap corresponding to the original location of the extracted constituent is deleted. However, this is clearly a stipulation. The failure of constraints to apply when there is no apparent extraction site is the problem, not the solution. The argument for deletion would be vastly more compelling if the BAE and sluicing examples in (7)–(12) were ungrammatical.

The facts of (7)–(12) are accounted for in SS by two simple assumptions: (i) there is no movement and (ii) there is no deletion. This of course raises the question of why the fragment appears to display the selectional properties that it would have if it were a constituent of a fuller structure. The fragment is not part of a larger syntactic structure, but it behaves as though it is. And it is interpreted as though it occupies a particular position in a syntactic structure, even though it doesn't. Culicover and Jackendoff 2005 call this phenomenon Indirect Licensing.

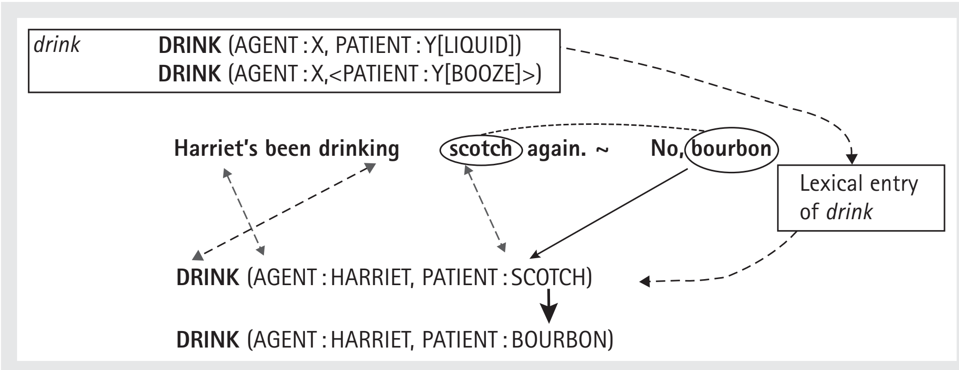


Figure 30.1. Interpretation of *bourbon* with respect to antecedent in (13)

Here are illustrations for a couple of simple examples. First, consider the following dialogue.

(13) A: Harriet has been drinking scotch again.
B: No, bourbon.

Here, *scotch* is the direct object of *drink* in the antecedent. *Bourbon* is an NP and therefore can match *scotch*. The interpretation of *bourbon* is substituted for that of *scotch* in the meaning of the antecedent to give the contextualized meaning of the fragment.

Next, consider the dialogue in (1). The lexical entry for *drink* indicates that the relation **DRINK** selects a PATIENT with the feature [LIQUID] when the argument is overt, and [BOOZE] when the argument is not expressed syntactically (indicated by the angled brackets in the second entry). In this latter case, an NP fragment is matched not with a constituent of the antecedent but with the selectional possibilities in the lexical entry.

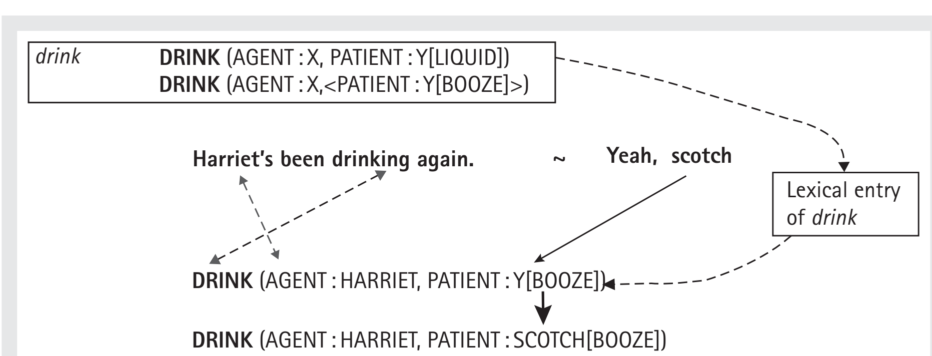


Figure 30.2. Interpretation of fragment "Yeah, scotch" for example (1)

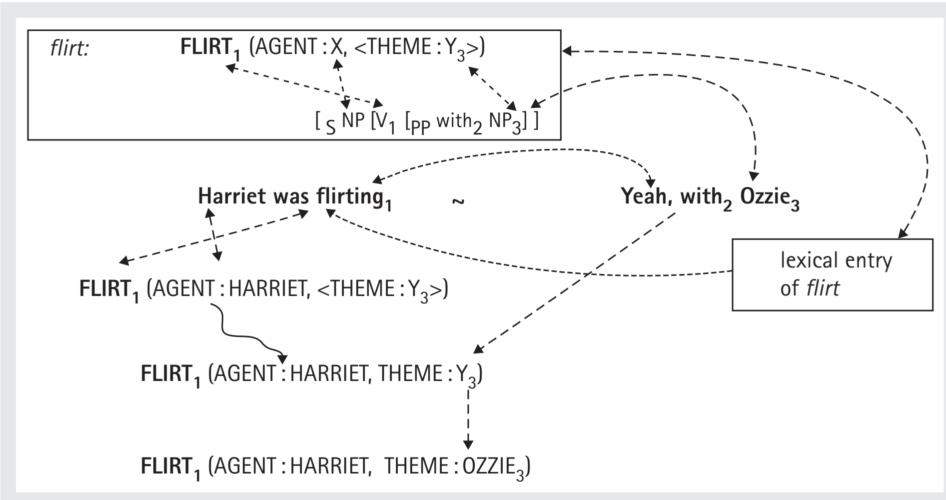


Figure 30.3. Interpretation of fragment *with Ozzie* in (4a)

Scotch, being an NP, has syntactic features that allow it to function as the direct object of *drink*. Consequently, the interpretation of *scotch* may be slotted into the PATIENT argument of **DRINK**. This produces the representation as shown.

Consider the well-formed dialogue in (4a). The verb *flirt* selects [$_{PP}$ *with NP*]. Hence the fragment *with Ozzie* satisfies this selectional requirement and corresponds to the THEME of **FLIRT**.

Finally, note that there are cases where the interpretation of the fragment must be sensitive to the pragmatics associated with the interpretation of the antecedent, e.g.,

(14) a. A: Ozzie said that Harriet's been drinking again.
B: Yeah, scotch.
[= 'Harriet's been drinking scotch again' or 'Ozzie said that Harriet's been drinking scotch again']

b. A: Ozzie mistakenly believes that Harriet's been drinking again.
B: Yeah, scotch.
[= 'Ozzie mistakenly believes that Harriet has been drinking scotch again'; ≠ 'Harriet has been drinking scotch again']

c. A: Ozzie doubts that Harriet has been drinking again.
B: * Yeah, scotch.
[No interpretation]

Thus an analysis in terms of deletion is in some cases wrong (because of the impossible extractions), in some cases unnecessary, and in some cases insufficient.

In SS a fragment is indirectly licensed through its match with the antecedent, and interpreted on the basis of the interpretation of the antecedent.

30.3 Simpler Syntax on control and raising

30.3.1 Control

Cases such as the following have been the primary focus of the study of control in MGG. (The control relation is marked with a subscript on the controlled predicate.)

(15) a. John$_i$ wanted to $_i$take care of himself$_i$.
b. Mary$_j$ convinced Susan$_i$ to $_i$take care of herself$_{i/*j}$.

Following IU, the complements are sentential, and the subjects are invisible pro-forms, typically represented as PRO, coindexed with the antecedent.

(16) a. John$_i$ wanted [PRO$_i$ to take care of himself$_i$]
b. Mary$_j$ convinced Susan$_i$ [PRO$_{i/*j}$ to take care of herself$_{i/*j}$]

These cases of control are characterized by the fact that the controller of the nonfinite clause must be an argument of the higher clause. It cannot be more distant, nor can the nonfinite clause have a generic controller.

(17) a. Susan$_i$ said that John$_j$ wanted to $_{*i/j}$take care of himself$_j$/*herself$_i$
b. *John wanted to $_{gen}$take care of oneself.

The observation that *Mary* cannot be understood as the subject of *take care*... in (15b) has been the principal empirical evidence that the relationship between PRO and the antecedent is a syntactic one: in both cases illustrated here, and in many others, the antecedent is the lowest NP in the tree that c-commands PRO. Control is thus assumed to be subject to a Minimal Distance Principle (MDP; cf. Rosenbaum 1967), stated over syntactic structure.

However, Rosenbaum 1967 noted an important counterexample to the MDP.

(18) Mary$_j$ promised Susan$_i$ [PRO$_{i/*j}$ to take care of herself$_{*i/j}$]

Here, the antecedent of PRO is the more distant NP, not the closest. Thus, it has been known for more than 40 years that the MDP is inadequate.

The empirical evidence, including (18), shows that the relationship between the antecedent and the controlled argument is not a syntactic one but a semantic one. Examples can be constructed in which the syntactic relations are varied but the semantic relations are held constant. When the semantic relations are held constant, the control relations are unchanged.

One particularly clear set of cases involves obligation. An order imposes an obligation on the Patient, while a promise imposes an obligation on the promiser. The obligation is necessarily with respect to one's own action—it is not possible to have an obligation to perform someone else's action. In the following examples, the locus of the obligation coincides with the antecedent of control.

(19) a. Bill ordered Fred$_i$ [to $_i$leave immediately]
b. Fred$_i$'s order from Bill [to $_i$leave immediately]
c. the order from Bill to Fred$_i$ [to $_i$leave immediately]
d. Bill gave Fred$_i$ the order [to $_i$leave immediately]
e. Fred$_i$ received Bill's order [to $_i$leave immediately]

(20) a. Bill$_i$ promised Fred [to $_i$leave immediately]
b. Fred's promise from Bill$_i$ [to $_i$leave immediately]
c. the promise from Bill$_i$ to Fred [to $_i$leave immediately]
d. Bill$_i$ gave Fred a promise [to $_i$leave immediately]
e. Fred received Bill$_i$'s promise [to $_i$leave immediately]

SS deals with control by making two assumptions: (i) there is no PRO, and (ii) control is semantic, not syntactic (hence no MDP). That such an approach is on the right track is shown further by the fact that the control relation in general is far more complex than the simple cases on which the MGG account is based. There is in fact long distance control (21a), split antecedent control (21b), generic control (21c), and speaker/hearer control (21d).[2]

(21) a. John$_i$ talked to Sarah$_j$ about $_{i/j}$taking better care of himself$_i$/herself$_j$.
b. John$_i$ talked to Sarah$_j$ about $_{i+j}$taking better care of themselves$_{i+j}$
c. John$_i$ talked to Sarah$_j$ about $_{gen}$taking better care of oneself$_{gen}$.
d. Undressing myself/yourself/ourselves [=you and me] in public could cause a scandal.

And, as noted by Postal 1969, there doesn't have to be an NP antecedent of control at all.

(22) a. another American attempt to dominate the Middle East
b. the Anglo-French agreement to respect each other's territorial claims

(23) a. Any similar attempt to leave will be severely punished.
b. Yesterday's orders to leave have been canceled.
c. How about taking a swim together?
d. How about giving each other a massage?

[2] A range of control possibilities were originally noted by Grinder 1970, Cantrall 1974, Williams 1985, Ladusaw and Dowty 1988, and Sag and Pollard 1991.

It thus appears that treating control as a semantic relation and not as a syntactic relation between an overt NP and an invisible NP is necessary.

The control relation in SS is one of binding at the level of semantic representation. For expository purposes I illustrate how this binding works using the relation of intention, which is somewhat more transparent than obligation. The essence of **INTENTION** is that the intender is necessarily the performer of the action. This is expressed in conceptual structure as follows, where X^{α} is the intender, and **R**[ACT] is the relation that denotes the act.

(24) **INTEND**(EXP : X^{α}, THEME : [**R**[ACT](AGENT : α, ...)])

INTEND is a component of the CS representations of a number of verbs, all of which have the same binding representation.

The correspondence rules for English state that if a bound argument α corresponds to the Subject GF (see section 30.1.1) and the relation **R** corresponds to a nonfinite predicate, the Subject GF is not realized phonologically. So the CS representation in (25a), which is paraphrased as "Sandy$^{\alpha}$ intends that Sandy$^{\alpha}$ will call", will correspond to (25b) if the predicate is *to call*.

(25) a. **INTEND**(EXP : SANDY$^{\alpha}$, THEME : [**CALL**(AGENT:α)])
b. Sandy intends to call.

The fact that Subject is not realized phonologically or syntactically when it is bound and the predicate is nonfinite gives rise to the appearance of "obligatory control" where the lexical entry of the verb requires that its nonfinite complement contain a bound argument, as in the case of **INTEND**. Such cases do not permit generic control, split antecedents, or speaker/hearer control, because the binding relation is stipulated as part of the lexical entry. This is what we find with verbs such as *try*, *intend*, and *promise*, which are so-called "obligatory subject control" verbs, and verbs such as *convince, persuade*, and *order*, which are so-called "obligatory object control" verbs. "Obligatory subject control" follows from the fact that in the lexical entry, for example of *intend*, the argument that corresponds to the subject of *intend* binds the Agent of the complement. Similarly for "obligatory object control", where the argument that corresponds to the object, for example the Patient of *persuade*, binds the Agent of the complement.

When there is a lexically defined binding relation, control will be obligatory, and it will be local. Locality follows from the fact that the only CS arguments that can be expressed in the lexical entry are the arguments of the higher relation, e.g., **INTEND**, and the arguments of the complement. Hence we expect to see the binding relation and the locality relation even when a language expresses control using "pro-drop" in a finite clause, as is the case in Greek and other Balkan languages (Joseph 1994; Landau 2000).

When there is no lexically specified control, it is still possible for a complement to correspond to a nonfinite predicate, and for the argument that corresponds to the subject to be bound. When this happens we get control, but since it is not lexically specified control, it can be free, subject to whatever other semantic or pragmatic restrictions might be in effect.[3] In this way we get cases of split antecedent control, generic control, and speaker/hearer control as illustrated by the examples (21).

We see that the SS approach to control in terms of CS binding is sufficient to account for obligatory control and less restricted types of control. The more restricted type of control is sensitive to semantic conditions associated with particular lexical relations, and is local for the same reason. In the absence of such semantic conditions, control is freer.

30.3.2 Raising

Let us turn to raising. There are traditionally two types of raising, raising to subject as in (26a) and raising to object as in (26b). (I use the traditional terminology for reference purposes only.)

(26) a. Susan seems to be proud of herself.
b. I believe Susan to be proud of herself.

Each has a full sentential paraphrase.

(27) a. It seems that Susan is proud of herself.
b. I believe that Susan is proud of herself.

Following IU, MGG treatments of raising to subject posit a movement from the subject position of the complement to the non-thematic subject position of *seem*.

(28) Susan$_i$ seems [t_i to be proud of herself]

In the case of raising to object, the subject of the complement functions as though it is the object of the higher verb—cf. *I believe her to be proud of herself, Susan is believed to be proud of herself*. In early treatments this NP became the direct object of the verb. In GB (Government and Binding Theory) it remained in situ and was governed by the verb and therefore (exceptionally) case-marked, and in recent approaches it moves, perhaps invisibly, into the specifier position of a functional head that licenses case. See Davies and Dubinsky 2004 for a review.

Raising as a movement is problematic because it has to be constrained. Example (29) shows that only the subject of a complement can be raised.

(29) a. *Robin$_i$ seemed for Sandy to dislike t_i.
[=It seemed Robin disliked Sandy.]

[3] Jackendoff and Culicover 2003 distinguish several types of control, with somewhat different properties.

b. *I believe Robin$_i$ for Sandy to dislike t_i.
[=I believe that Sandy dislikes Robin.]

The examples in (30) show that raising to subject must apply locally. (Similar examples for raising to object can be constructed but are more complex—I will not give them here.)

(30) a. *Robin$_i$ seemed [that it would rain] [without t_i getting dressed]
[Raising of *Robin* out of adjunct *without Robin getting dressed*]
b. *Robin$_i$ seemed [that it would be fun [t_i to yodel all by herself]]
[Long distance raising of *Robin* out of complement of complement]
c. * Robin$_i$ seemed that [it was obvious [that [t_i to yodel all by herself] would not be a whole lot of fun]]
[Long distance raising of subject of complement of complement]

Comparable examples show that control is not constrained by locality.

(31) a. Robin$_i$ can yodel without $_i$giggling. [Control into adjunct]
b. Robin$_i$ thought [that it would be fun $_i$[to yodel all by herself$_i$]]]
[Long distance control into complement of complement]
c. Robin$_i$ thought that [it was obvious [that $_i$[to yodel all by herself$_i$] would not be a whole lot of fun]]
[Long distance control into subject of complement of complement]

It is of course no problem to stipulate constraints on the movements used to derive raising so that the ungrammatical examples are not generated. In fact, much of the Conditions framework of Chomsky 1973 and GB theory of Chomsky 1981 and Chomsky 1986*b* is built on the formulation of such constraints. SS deals with the problem differently, by making the following two assumptions: (i) there is no raising, and (ii) the thematic roles of "raised" arguments are assigned through the linking of GFs.

The second assumption reflects the fact that the GFs are used in SS to capture certain relations that are captured using movement in MGG. Unlike movement, the use of GFs is inherently local. Figure 30.4 shows a simple correspondence between syntactic structure, GFs, and conceptual structure (CS).

Whether or not there is an Object GF depends on the argument structure associated with the CS relation, and idiosyncratic properties of the corresponding verb. The verb *love* has an Object, the verb *fall* does not, and *look for* expresses the Theme argument obliquely, and not as an Object. But regardless of the verb, there is a Subject by default. Subject corresponds to a full NP when it expresses a thematic argument, to a dummy (as in the case of extraposition—*it is obvious that S*), or to null in the case of non-finite complements.

In the case of complementation, one CS relation is embedded as an argument of another. The following illustrates for the relations **SEEM** and **BELIEVE** and the CS representation corresponding to *Sandy likes chocolate*.

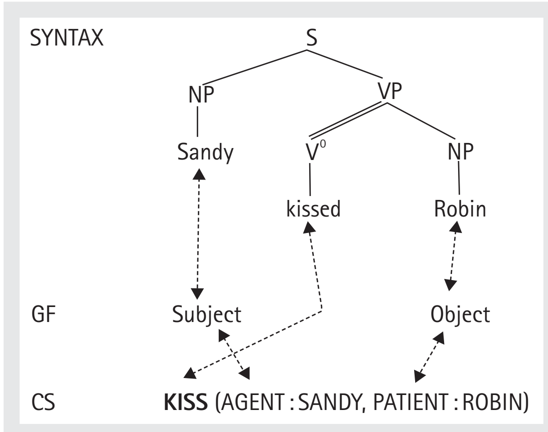


Figure 30.4. Simple correspondence showing Syntax, CS, and GFs

(32) **SEEM**(**LIKE**(EXP : SANDY, THEME : CHOCOLATE))

(33) **BELIEVE**(EXP : ROBIN, THEME : [**LIKE**(EXP : SANDY, THEME : CHOCOLATE)])

The normal linking of the Experiencer of **LIKE** maps it to the Subject GF in the subordinate clause. But this clause, being infinitival, lacks a syntactic NP preceding the VP. The relation called "raising" links the unassigned Subject GF to a GF in the upper clause, where it can then be realized as an NP. Figure 30.5 and Figure 30.6 illustrate, respectively. Note that in the case of raising to subject, the Subject GF of the complement, which is annotated as Subject$_1$, is linked to the

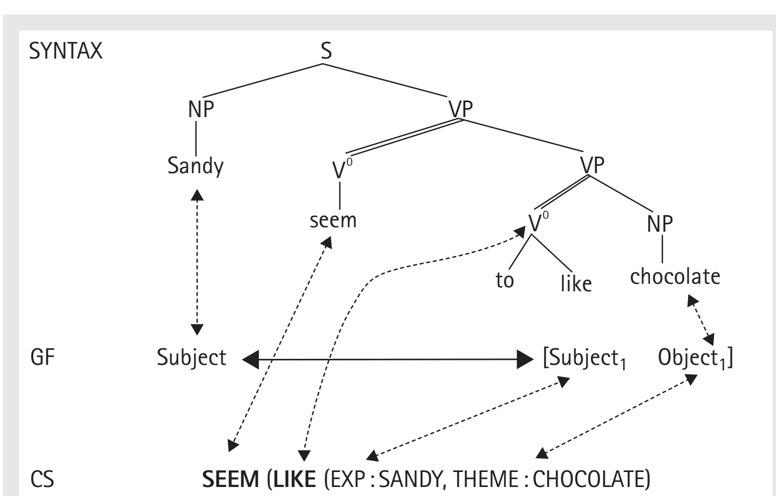


Figure 30.5. Raising to subject in SS

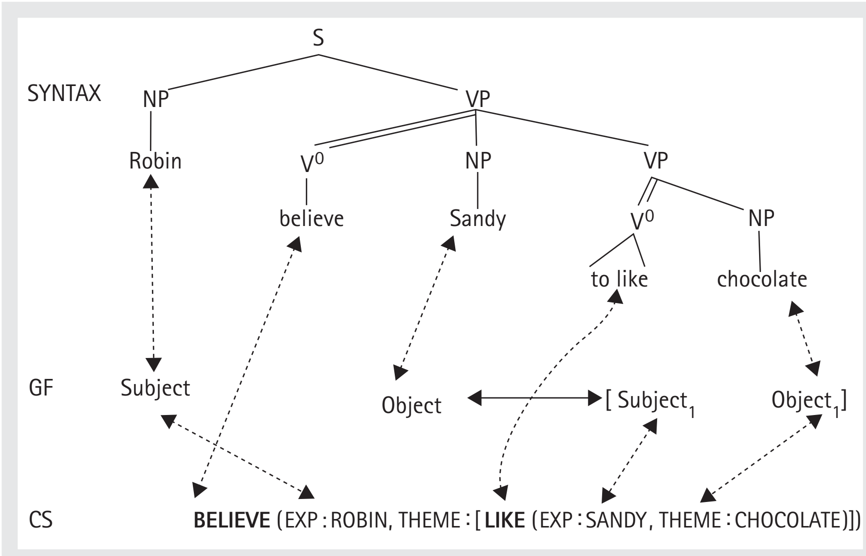


Figure 30.6. Raising to object in SS

Subject of *seem*, while in the case of raising to object, it is linked to the Object of *believe*.

Characterizing raising in terms of a correspondence between the Subject GF of the complement and an argument of the higher clause has the consequence that the relation will always be local. Moreover, an argument of an adjunct subordinate clause cannot participate in this relation, because an adjunct subordinate clause is not selected by a verb. Thus its subject is not accessible in the lexical entry of the verb.

The fact that a particular predicate permits this type of relation is one that has to be stated explicitly in the lexical entry for that verb regardless of the theory, since not all verbs permit raising to subject and raising to object, even when the semantics suggest that it is possible. For example, *unlikely* allows raising to subject, while *improbable* does not.

(34) Sandy is $\left\{\begin{matrix}\text{unlikely}\\ \text{*improbable}\end{matrix}\right\}$ to call.

Similarly, *expect* allows raising to object, but *anticipate* does not.

(35) Robin $\left\{\begin{matrix}\text{expects}\\ \text{*anticipates}\end{matrix}\right\}$ Sandy to call.

The SS account takes advantage of lexical information that is independently required to capture the behavior of these constructions. Movement is unnecessary.

The switch to a SS account of raising in terms of GFs is able to dispense with locality constraints on movement. In the next section, I discuss how it is possible to dispense with movement in general, while capturing the generalizations.

30.4 Simpler Syntax on movement

30.4.1 Why movement?

The motivation for movement in MGG is the observation that in certain constructions, constituents that are not in their canonical positions have the semantic and grammatical functions that are associated with these positions. Uniformity considerations lead to the conclusion that the relationships should be expressed in terms of movement. I highlight the two basic cases in which the application of the uniformity methodology leads to this conclusion.

In certain cases, termed A movement, only the semantic functions are associated with the non-canonical positions. For instance, in the passive construction in English, the subject NP bears the thematic role normally associated with the direct object.

(36) a. Sandy kissed Robin.
AGENT THEME

b. Robin was kissed by Sandy
THEME AGENT

Assuming uniformity (IU), the thematic roles must be associated uniformly with a syntactic configuration. The simplest hypothesis is that the Theme role is associated with the direct object configuration. It follows that *Robin* in (36b) must be a direct object. Since this NP is demonstrably the subject of (36b), there must be two syntactic representations, one in which *Robin* is direct object and one in which *Robin* is subject, and *Robin* must move from object to subject.

The other type of movement arises when the constituent has not only the semantic but the grammatical functions associated with the canonical position. These are the A′ movements, exemplified by wh-questions in English.

(37) Who did Sandy kiss?
THEME AGENT

Standard tests show that *who* is the direct object and *Sandy* is the subject. Again, uniformity considerations lead to the hypothesis that there are two syntactic representations, one in which *who* is in direct object position following the verb, and one in which it is in clause-initial position.

SS seeks to account for the argument alternations that motivate A movement simply in terms of correspondences between CS and syntax, mediated by GFs. SS shows that these alternations can be accounted for with no loss of generalization. Hence A movement is unnecessary. Moreover, there are argument alternations that must be accounted for in terms of independent lexically specified correspondences. Hence A movement is insufficient. I develop this argument in section 30.4.2.

SS seeks to account for the phenomena observed in A′ constructions by linking the constituent in A′ position to its canonical position without movement. SS shows how it is possible to define "chains" without movement, using CS representations to mediate the relationship between the constituent in A′ position and the canonical position. Hence A′ movement is unnecessary. Moreover, there are chains that cannot be derived in terms of movement that are mediated by CS representations in just the way that A′ chains are. Hence A′ movement is insufficient. I develop this argument in section 30.4.3.

30.4.2 Argument alternations in Simpler Syntax

Accounts of argument alternations without A movement have been fully developed in HPSG and LFG. The SS account is a distillation of these accounts, formulated in terms of the correspondences between CS arguments and GFs, and the correspondences between GFs and syntactic configurations. In this section I summarize how SS expresses the active/passive relation, the antipassive, and the applicative relation in these terms.

(*i*) *active/passive*

There are two fundamental GF/syntax correspondence rules in English, one for Subject and one for Object.

Which argument corresponds to Subject and which to Object is a matter of linking. Some verbs specify how their arguments are expressed syntactically, while others follow the default linking. The default linking is given in (38).

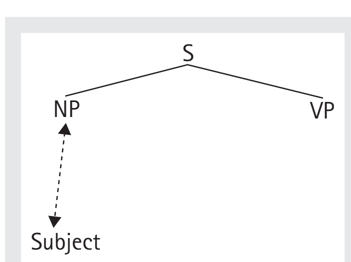


Figure 30.7. Subject correspondence rule for English

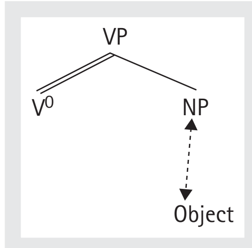


Figure 30.8. Object correspondence rule for English

(38) *Default linking*

$$\left\{\begin{array}{c}\text{Agent/Experiencer}\\ \Downarrow\\ \text{Theme/Patient}\end{array}\right\} \leftrightarrow \left\{\begin{array}{c}\text{Subject}\\ \Downarrow\\ \text{Object}\end{array}\right\}$$

What this says is that in the default case, if there is an Agent/Experiencer, it corresponds to Subject, and then if there is a Theme/Patient that corresponds to Object. But if there is no Agent/Experiencer, then Theme/Patient corresponds to Subject.[4]

A simple case in which the Subject and Object correspondences and the default linking apply is given in Figure 30.4 above. For such cases, it is sufficient to state the CS representation associated with a verb as part of its lexical entry; the defaults take care of the correspondences when the verb is part of a sentence. We use a standard attribute-value matrix (AVM) to specify the lexical information.

(39) *kiss*

$$\begin{bmatrix}\text{SYNTAX} & [\text{CATEGORY} \quad \text{V}]\\ \text{CS} & \textbf{KISS}(\text{AGENT : X, PATIENT : Y})\end{bmatrix}$$

There are two types of cases in English that are somewhat more complex than that of *kiss*. One involves verbs whose linking does not fall under the default (38). One verb that specifies its linking is *receive*, meaning "come to possess". The Location of possession is linked to Subject, and Theme is linked to Object. In this case we have to complicate the lexical representation by making reference to the Subject GF.

(40) *receive*

$$\begin{bmatrix}\text{SYNTAX} & [\text{CATEGORY} \quad \text{V}]\\ \text{GF} & \text{Subject}[1]\\ \text{CS} & \textbf{BECOME}(\textbf{HAVE}(\text{LOCATION : X}_1\text{, THEME : Y}))\end{bmatrix}$$

[4] It may be possible to replace this default linking with the considerably more nuanced linking of Dowty 1991, though see Jackendoff 2007: 203 for a critique of Dowty's analysis.

The other type of more complex case is where one or more of the CS arguments do not correspond to Object but to a prepositional phrase. We call this an oblique object. Typically, the preposition is specified in the lexical entry of the verb. For example, *look for* has the same CS representation as *seek*, but represents the Theme as the complement of the preposition *for*. The AVM in (41) shows the PP *for NP* as the complement of *look*, and coindexes the complement of *for* with the Theme argument.

(41) *look (for)*

$$\begin{bmatrix} \text{SYNTAX} & \begin{bmatrix} \text{CATEGORY} & \text{V} \\ \text{COMPS} & [_{\text{PP}}\ \textit{for}\ \text{NP}_1] \end{bmatrix} \\ \text{CS} & \textbf{SEEK}(\text{AGENT : X, THEME : Y}_1) \end{bmatrix}$$

Let us consider now the passive. Chomsky 1957 pointed out the following properties of the passive construction in English.

(42) a. The passive participle following a form of *to be* occurs only with a transitive verb.
b. V in the passive cannot be followed by a direct object.[5]
c. An agentive *by*-phrase can occur only if the sentence is passive.
d. The selectional restrictions on subject and object of the active are mirrored in the selectional restrictions on the *by*-phrase and subject of the passive, respectively.

On the basis of the IU and (42d), Chomsky concludes that the passive is derived transformationally from the structure that underlies the active. But Emonds 1970 showed that the syntactic structure of the passive is a special case of the canonical structure of English, with an NP in canonical subject position, the auxiliary verb *be* preceding what appears to be an adjectival predicate, and a PP in VP final position. This "structure-preserving" character of the passive has motivated a series of proposals over the years that treats the passive as basic, with an alternate linking to thematic structure; see Bresnan (1978; 1982*b*); Brame (1978); Hudson (1984); Gazdar et al. (1985); Pollard and Sag (1994), and many others.

The key syntactic property of the passive is that the link between the highest argument in the linking hierarchy (38) is suppressed, and the next highest argument is then linked to Subject. (In Relational grammar this relation is expressed by shifting the subject to an oblique function (Perlmutter and Postal 1983), while in GB theory it is represented by "absorbing" the θ-role of the subject into the passive morphology (Baker et al. 1989)). Since the CS representation is unchanged, observation (42d) follows immediately. Observation (42c) must be stipulated in any framework as a special property of the passive construction. Observations (42a,b)

5 With caveats for examples like *Sheila was sent flowers*. In this case, it is the indirect object that does not follow the verb.

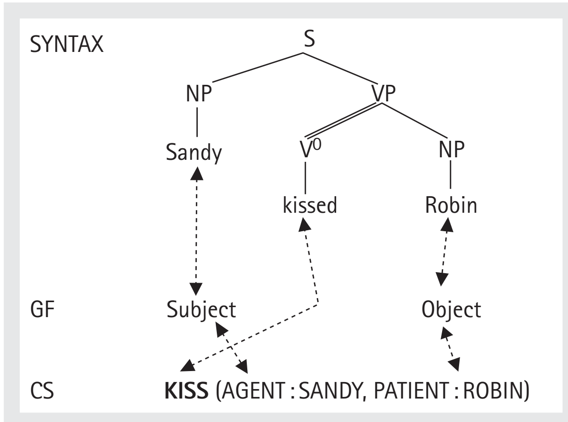


Figure 30.9. Correspondences for *Sandy kissed Robin*

are captured by assigning to the auxiliary verb *be$_{passive}$* the lexical representation in (43).

(43) *be* (passive)

$$\left[\text{SYNTAX} \begin{bmatrix} \text{CATEGORY} & \text{V} \\ \text{COMPS} & \begin{bmatrix} \text{VP} \\ \text{PASSIVE} \end{bmatrix} \end{bmatrix} \right]$$

This verb selects a passive VP whose head is marked with passive morphology.

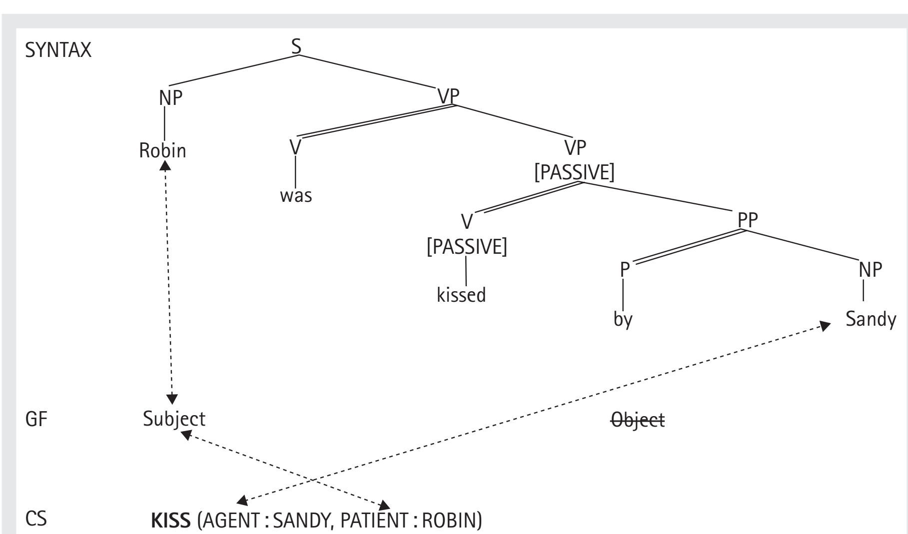


Figure 30.10. Correspondences for *Robin was kissed by Sandy*

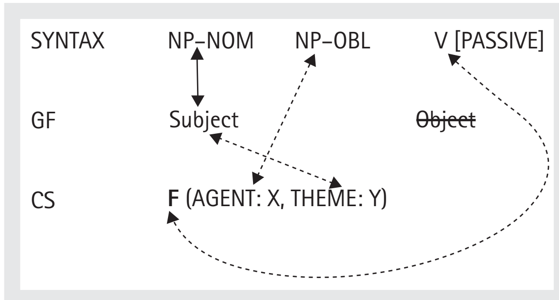


Figure 30.11. Schema of passive correspondence in nominative-accusative language

In SS, the sentences in (36) have the correspondences in Figure 30.9 and Figure 30.10. No movement is involved in the passive, there is simply a different correspondence between the CS arguments and the syntactic arguments.

Finally, in a nominative-accusative language, such as Russian, the GFs are realized by morphological case, but the suppression of the subject functions as it does in English; see Figure 30.11. The passive in such a language has the same realignment of the CS arguments with the GFs, but the GFs are realized differently than in a language such as English.

A natural idea at this juncture is that the characterization of the passive in terms of an alternative correspondence for the grammatical functions is a notational variant of the movement analysis. However, SS, following the insights of Relational Grammar, takes the key property of the passive to be how the canonical subject is expressed—it is either suppressed, or realized as an oblique argument. The correspondence of the canonical object with the Subject GF is not movement but follows from the default linking.

That this is the right way to think of the passive is shown by the fact that there are other options in natural language that may arise when the highest argument in the hierarchy is suppressed. In some languages, such as German, when the canonical subject is suppressed and the verb is intransitive, the subject position may be realized as an expletive.

(44) *German*

a. Es wurde getanzt.
it be(come).PAST.3SG dance-PAST.PART.
'There was dancing.'

b. Es wurde viel gelacht.
it be(come).PAST.3SG much laugh-PAST.PART.
'There was a lot of laughing.'

It is sufficient to say that German allows this type of impersonal passive because it requires the subject position to be filled with *es* when the subject position is not filled with a full NP.[6] In English, the expletive subject option is available only for a restricted set of constructions that does not include passive. In contrast, in MGG, the movement of the object is a key component of the analysis of the passive.[7]

Similarly, impersonal passives are found in languages such as Italian, marked with the reflexive.

(45) *Italian*

a. In Italia tutti mangiano spaghetti.
in Italy everyone eat-PRES-3PL spaghetti-PL
'In Italy everyone eats spaghetti.'

b. In Italia si mangia spaghetti (*per tutti).
in Italy SI eat-PRES-3SG spaghetti-PL (by everyone)
'In Italy spaghetti is eaten.'

Note that in (45b) the singular verb does not agree with the plural *spaghetti*, showing that *spaghetti* is the object and not the subject. In this case, the possibility of an unexpressed "pro" subject with the reflexive allows the canonical object to link to the Object GF. The effect is the same as that of the passive construction in English, but with a different syntactic realization of the arguments.

Finally, in Manggarai, a language in the Bima-Sumba subgroup of Indonesian languages, there is no passive morphology, but there is marking of the Agent as an oblique argument.

(46) *Manggarai*

a. Aku cero latung-k
1SG fry corn-1SG
'I fry/am frying corn.'

b. [Latung hitu] cero l-aku-i
corn that fry by-1SG-3SG
'The corn is (being) fried by me.'

(Arka and Kosmas 2005)

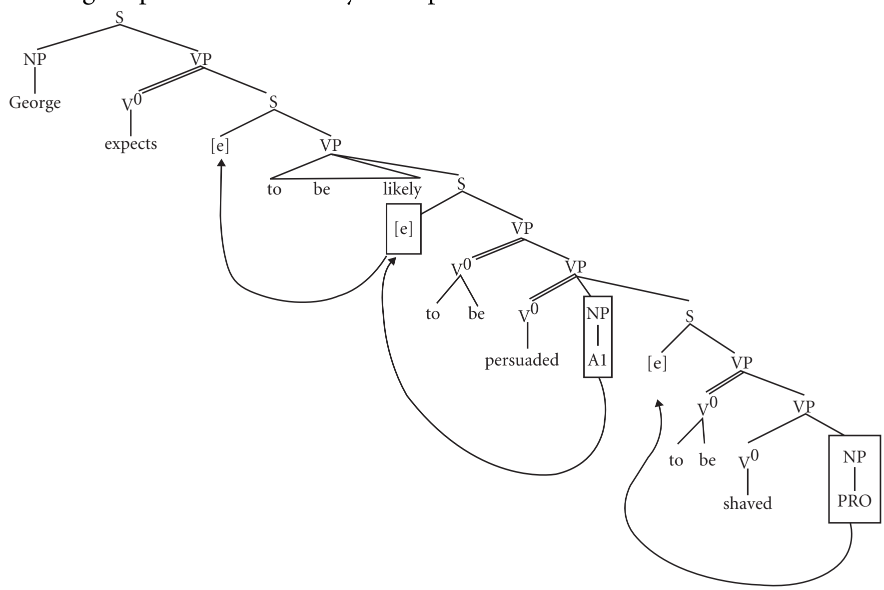
Manggarai is a language in which grammatical roles are marked by word order, so it can be argued that *latung hitu* "the corn" in (46b) is the subject, hence *l-aku-i* "by me" is an oblique argument. Since the verb is not marked with passive morphology, it is hard to argue that it is comparable to the intransitive adjectival or participial that occurs in English. In this case it is plausible that there is simply an alternative way of expressing the canonical object that arises from the alternate correspondence for the canonical subject.

[6] I refer to the position of *es* as subject position, but it may well be the obligatory main clause topic position. Cf. *Gestern wurde getanzt* "Yesterday there was dancing", **Gestern wurde es getanzt*.

[7] Chomsky 1981.

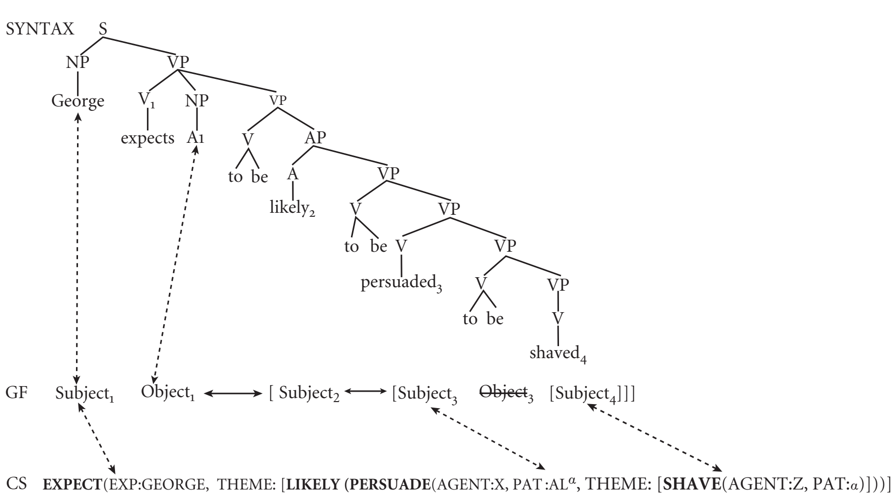
The important point is that these constructions are other ways of suppressing the highest CS argument without syntactic movement and without passive morphology. This can be shown even for complex cases as well. One of the most impressive achievements of MGG was to show that chains of simple movements of an NP from object to subject, and then raising of that derived subject, could produce complex grammatical patterns. An example is given in (47).

(47) George expects Al to be likely to be persuaded to be shaved.



Such cases are not a problem for a non-movement approach such as SS, because there may be multiple correspondences between GFs.

(48)



There are two passives, one "raising to subject" predicate and one "raising to object" predicate. Because *shaved* is a passive, α, which is the Patient of **SHAVE**, corresponds to the Subject GF of the lowest predicate. α is controlled, hence bound, by AL^{α}, which is the Patient of **PERSUADE**. Since *persuaded* is passive, AL^{α} corresponds to the Subject of the *persuaded* clause. Since *persuaded* is the infinitival complement of *to be likely*, which is a "raising" predicate, its Subject corresponds to the Subject of *to be likely*. But this clause is the infinitival complement of *expects*, which is a "raising to object" verb. So this Subject corresponds to the Object of *expects*.

In summary, SS accounts for the passive by making two simple assumptions: (i) there is no movement, and (ii) the highest argument is suppressed or corresponds to an oblique. Dispensing with movement means that there is no need to develop a theory of case licensing to trigger movement, a theory that is particularly problematic for languages that lack morphological case. This in turn eliminates some of the motivation for invisible functional heads whose purpose is to license case (or other features, such as EPP) and trigger movement.

SS needs to assume the default linking hierarchy, primitive GFs, and alternate correspondences between CS arguments and GFs. In the remainder of this section I briefly note a number of other phenomena that suggest that these devices are independently required, and therefore do not constitute an additional theoretical cost.

(ii) antipassive

The passive permits suppression of the highest CS argument or makes it an oblique argument. In the antipassive, the second argument is suppressed or expressed as an oblique. Antipassive is sometimes called detransitivization because it makes a transitive into an intransitive while holding the subject constant.

Antipassives are seen typically in ergative languages. In the transitive, the subject is marked with the ergative case, while the object is marked with the absolutive case. But in the antipassive, which is in transitive, the subject is marked in the absolutive, and the object is realized as an oblique. The data from Chukchee in (49) illustrate.

(49) *Chukchee*

a. ʔaaček-a kimit-ən ne-nlʔetet- ən
youth-ERG load-ABS 3PL-A-carry-3SGP
'The young men carried away the load.'

b. ʔaaček-at ine-nlʔetet-gʔe-t kimit-e
youth-ABS ANTIPASS-carry-3PLS load-INSTR
'The young men carried away a load.'

(Kozinsky et al. 1988, cited by Kroeger 2004: 293)

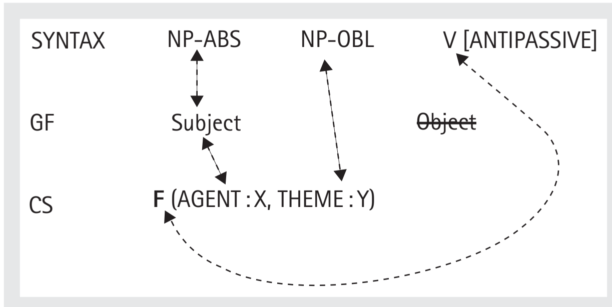


Figure 30.12. Schema for antipassive correspondence in ergative-absolutive language

The suffix on the verb agrees with the phrase marked with the absolutive case. In the antipassive example (49b), the instrumental is optional, just like the *by*-phrase is in the English passive.

It is somewhat less natural to envision a movement account of the antipassive because there is no advancement of the direct object to subject.[8] The SS approach treats it in a way that is parallel to the passive—compare Figure 30.12 with Figure 30.11.

The antipassive is a regular, morphologically marked construction. In this respect it is similar to the passive. In English there are alternations that resemble the antipassive, but, crucially, they are lexical alternations. Some examples are given in (50).

(50) a. Kim was eating cereal.
Kim was eating. (i.e., eating something)
b. Sandy was drinking beer.
Sandy was drinking. (i.e., drinking something)
c. Chris was reading the newspaper.
Chris was reading. (i.e., reading something)
d. Terry was cooking dinner.
Terry was cooking. (i.e., cooking something)
e. Marty was writing a letter.
Marty was writing. (i.e., writing something)
f. Leslie was chewing the gum.
Leslie was chewing (i.e., chewing something)

There are many verbs in English that do not permit this option.

[8] It is not possible to rule out a movement account. See Bobaljik and Branigan 2007 for one such proposal.

(51) Kim was {
*burning. (i.e., burning something)
*covering. (i.e., covering something)
*opening. (i.e., opening something)
*seeing. (i.e., seeing something)
*fixing. (i.e., fixing something)
*consuming (i.e., consuming something)
*writing {down / out / up} (i.e., writing {down / out / up} something)
}

In addition, there are some lexical alternations involving a single verb where the logical object argument is realized either as a direct object or an oblique object, with a subtle meaning difference.

(52) a. The dog was chewing (on) the book
b. We clutched (at) the rope.
c. The chimp climbed (up) the tree.
d. I was reading (in) the newspaper.

Such lexical alternations resemble what is found in the antipassive, where similar subtle meaning differences are often found.

In order to account for the English alternations, we must posit alternative ways of syntactically realizing the CS arguments. Given that this is required for such lexical alternations, there is no reason why the same device should not be used for regular morphologically mediated relations like the antipassive. There is no evidence that the English alternations are related by movement. Therefore there is no reason to believe that movement is implicated in the derivation of the antipassive.[9]

(iii) applicative

Because of space limitations I will simply note that the type of case that is made above for the antipassive can be made as well for the applicative. The applicative applies to the direct object, which is either suppressed or expressed as an oblique. At the same time, a CS argument that would canonically be expressed as an oblique is expressed as a direct object. The pattern is similar to the passive, but does not implicate subjects. The following examples from Chichewa illustrate the relation.[10]

(53) *Chichewa*

a. Mbidzi	zi-na-perek-a	msampha	kwa	nkhandwe.
zebras	SP-PAST-hand-ASP	trap	to	fox

'The zebras handed the trap to the fox.'

[9] It is of course possible to formulate an analysis of the English lexical alternations in terms of movement, but this is quite different from a demonstration that movement is necessary for the analysis.

[10] The gloss SP is a prefix that is used when there is a full NP subject.

b. Mbidzi zi-na-perek-er-a nkhandwe msampha.
zebras SP-PAST-hand-to-ASP fox trap
'The zebras handed the fox the trap.'

(54) a. Ndi-na-tumiz-a chipanda cha mowa kwa mfumu.
1SG.SUBJ-PAST-send-ASP calabash of beer to chief
'I sent a calabash of beer to the chief.'

b. Ndi-na-tumiz-ir-a mfumu chipanda cha mowa.
1SG.SUBJ-PAST-send-to-ASP chief calabash of beer
'I sent the chief a calabash of beer.'

(55) a. Fisi a-na-dul-a chingwe ndi mpeni.
Hyena SP-PAST-cut-ASP rope with knife
'The hyena cut the rope with a knife.'

b. Fisi a-na-dul-ir-a mpeni chingwe.
hyena SP-PAST-cut-with-ASP knife rope
'The hyena cut the rope with a knife.'

(56) a. Msangalatsi a-ku-yend-a ndi ndodo.
entertainer SP-PRES-walk-ASP with stick
'The entertainer is walking with a stick.'

b. Msangalatsi a-ku-yend-er-a ndodo.
entertainer SP-PRES-walk-with-ASP stick
'The entertainer is walking with a stick.'

(Baker 1988: 229, 230, 238, 260)

As in the case of the antipassive, there is a lexical relation in English, namely the dative alternation, that resembles the applicative. As is well known, the dative alternation does not apply to all verbs.

(57) Chris $\left\{\begin{array}{l}\text{gave}\\\text{showed}\\\text{loaned}\\\text{sold}\\\text{sent}\\\text{mailed}\\\text{took}\\\text{brought}\\\text{awarded}\\\text{assigned}\\\text{bequeathed}\end{array}\right\}$ $\left\{\begin{array}{l}\text{the money to Sandy}\\\text{Sandy the money}\end{array}\right\}$.

(58) Chris $\left\{\begin{array}{l}\text{built}\\\text{bought}\\\text{found}\end{array}\right\}$ $\left\{\begin{array}{l}\text{a house for Sandy}\\\text{Sandy a house}\end{array}\right\}$.

(59) Chris $\left\{\begin{matrix}\text{donated}\\ \text{presented}\\ \text{pushed}\\ \text{committed}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{the money to Sandy}\\ \text{*Sandy the money}\end{matrix}\right\}$.

(60) Chris $\left\{\begin{matrix}\text{constructed}\\ \text{purchased}\\ \text{created}\\ \text{invented}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{a house for Sandy}\\ \text{*Sandy a house}\end{matrix}\right\}$.

Again, it is technically possible to analyze the applicative and the dative alternation in terms of movement; see, for example, Larson 1988 for such an analysis of the English dative. For the lexical alternation, a movement analysis requires that each verb that participates in it have a diacritic feature that produces the desired output, according to the specific details of the analysis. I will not go into the details here; for a critique of Larson's proposal, see Jackendoff 1990*b*.

A more general point is that the alternation can be handled straightforwardly by specifying the possible correspondences of the CS arguments, along the lines of (61). Note the coindexing of the Goal argument with the object of the preposition *to* in (61a) and with the first Object NP in (61b).[11]

(61) a. *give₁*

$$\begin{bmatrix} \text{SYNTAX} & \begin{bmatrix} \text{CATEGORY} & \text{V} \\ \text{COMPS} & \text{NP}[_{\text{PP}}\ \textit{to}\ \text{NP}_1] \end{bmatrix} \\ \text{CS} & \textbf{GIVE}(\text{AGENT/SOURCE} : \text{X}, \text{THEME} : \text{Y}, \text{GOAL} : \text{Z}_1) \end{bmatrix}$$

b. *give₂*

$$\begin{bmatrix} \text{SYNTAX} & \begin{bmatrix} \text{CATEGORY} & \text{V} \\ \text{COMPS} & \text{NP}_1\ \text{NP} \end{bmatrix} \\ \text{CS} & \textbf{GIVE}(\text{AGENT/SOURCE} : \text{X}, \text{THEME} : \text{Y}, \text{GOAL} : \text{Z}_1) \end{bmatrix}$$

30.4.3 Chains in Simpler Syntax

Let us turn now to A′ chains. I assume without discussion that in an A′ construction there is a gap in the canonical position.

Applying IU to A′ constructions such as topicalization results in an analysis in which the gap is produced by movement. For example, in (62) the clause initial constituent functions as the direct object of the verb *like*.

[11] Larson motivated a movement analysis by the observation that the binding of anaphors in VP appears to reflect linear order, and not simply c-command as noted originally by Barss and Lasnik 1986. The movement analysis allows one to preserve c-command as a condition for binding, by permitting the binding relation to be stated over a structure that is subsequently transformed through movement. If the SS account of the dative alternation is correct, it has the consequence that c-command cannot be relevant for binding, and that GF and the relationships among CS arguments must form part of the account.

(62) Him_i, I like t_i.

There are two problems: (i) how to relate a constituent in an A′ position to the gap, and (ii) how to account for the gap when there is nothing in A′ position. In order for him_i to function in this way it must be linked to the canonical position through a chain. The parts of the chain are indicated by coindexing.

Regarding (i), movement must be constrained so that constituents only move higher in the syntactic structure; this is why (63a) is impossible. And constituents may move only to c-commanding positions; this is why (63b) is impossible.

(63) a. *t_i said [who_i that it was going to rain]
b. *[Robin who_i] said [that you were looking for t_i]

Regarding (ii), applying DU to A′ constructions such as English relative clauses results in an analysis in which there are invisible or deleted constituents in clause-initial position.

(64) a. the man who_i I like t_i
b. the man that $\cancel{\text{who}_i}$ I like t_i
c. the man $\cancel{\text{who}_i}$ I like t_i

SS considers as complications of syntactic theory (a) the conclusion that there is movement, (b) the need to represent a syntactic chain in terms of coindexing, (c) the consequence that movement has to be constrained, and (d) the fact that in some cases there is no overt constituent in A′ positions. These complications should be eliminated, if possible. The goal of eliminating complications is of course widely shared. In the Minimalist Program, for example, we see the attempt to eliminate coindexing and to derive the upward c-commanding character of chain formation from the basic architecture; see Chomsky 1995.

But the Minimalist Program does not eliminate movement per se in the formation of chains. And because it assumes DU, it does not avoid the consequence that there must be invisible constituents moving to A′ position. As far as I can tell, the primary motivation for holding onto movement as the way to create chains is that it allows for the possibility that derivations can be compared in terms of derivational economy. On this view, the more complex derivation is preempted by the less complex derivation. Chomsky explores the question of whether there is empirical support for this perspective but the evidence is less than conclusive.

SS deals with A′ chains by assuming that there is no movement and that gaps are freely introduced into syntactic representations. The precedents for such an approach in the literature go back at least to the 1970s—see for example Bresnan 1978; Brame 1978; Koster 1978. The problem of relating the A′ constituent to the gap is accomplished in SS by linking the A′ constituent to an operator in CS and the gap to a variable bound by this operator. The correspondence for clause-initial wh is given in Figure 30.13.

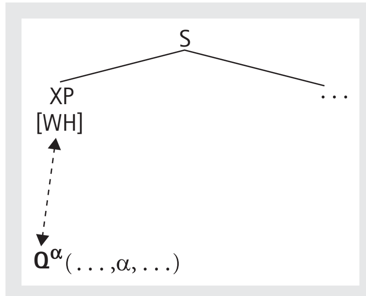


Figure 30.13. Correspondence for clause-initial wh

When there is a transitive verb and no overt post-verbal NP, NP is realized as a gap [*e*]. This gap corresponds to the variable α bound by the interrogative operator $\mathbf{Q}^{\alpha}$. A typical case is illustrated in Figure 30.14.

Because *what* corresponds to $\mathbf{Q}^{\alpha}$, which binds α, which corresponds to the gap, there is a chain between *what* and the gap that is mediated by the CS representation. The constraints on chains noted above do not have be stipulated as they do on a movement account. The wh-phrase is higher than the gap and c-commands it because the wh-phrase corresponds to a position on the left edge of the clause, while the gap corresponds to some argument position.

The characterization of a chain as a relation mediated by CS has the welcome property that there does not have to be a displaced constituent in the syntax in order for there to be a chain. All that is necessary is that a binding operator in

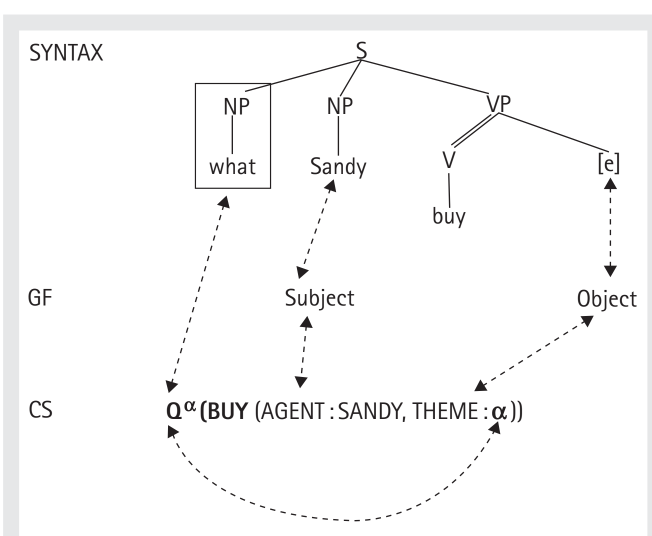


Figure 30.14. Correspondence for wh-question

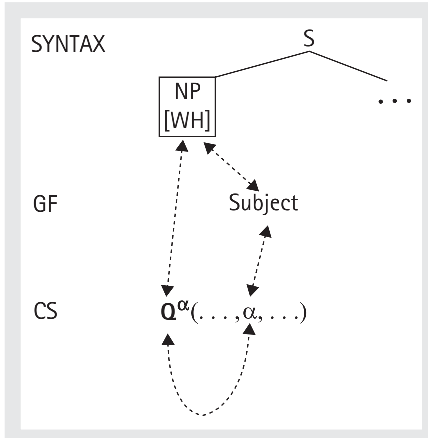


Figure 30.15. Subject wh correspondence

CS and the variable correspond to something in the syntactic representation. For example, wh-subjects satisfy both the wh correspondence of Figure 30.13 and the subject correspondence of Figure 30.7, and the chain does not involve a gap. In relative clauses there is a gap, but the head of the NP may correspond to the binder in CS. Similar correspondences will incorporate chains for constructions such as topicalization, infinitival relatives (with and without wh-phrases), *tough* movement, *too/enough*, the comparative correlative, and comparatives.

This brings us to the question of constraints on extraction. Chomsky (1977) pointed out that all A′ constructions obey the island constraints of Ross (1967) and Chomsky (1973). DU suggests that they all involve movement to an A′ position. If we abandon DU, as we do in SS, then we must attribute the behavior of A′ constructions not to movement but to the construction of chains whose tails are gaps. Koster (1978) showed how to do this in MGG, and there is a similar demonstration in HPSG by Pollard and Sag (1994). A possibility that I do not have space to go into here is that some extraction constraints arise out of the processing complexity of the configuration that contains the chain; see for example Hawkins (1994), Culicover (1999); Arnon et al. (in press); Phillips (2006).

As in the case of A movement, it might appear at first sight that a non-movement account of A′ constructions is a notational variant of a movement account. But it has been observed many times in the literature that there are A′ chains that do not contain gaps, and these cannot be derived through movement.

Let us suppose that the representation of a sentence with a discourse topic is roughly that of (65b), where for clarity CS is distinguished from Information Structure (IS). The topic corresponds to a **TOPIC** operator in IS.

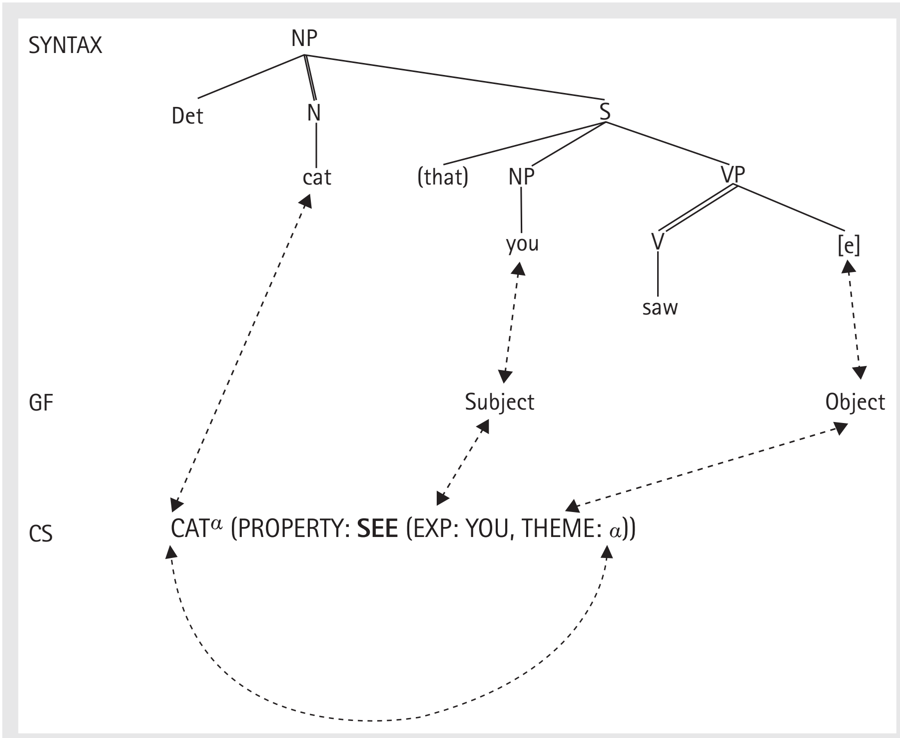
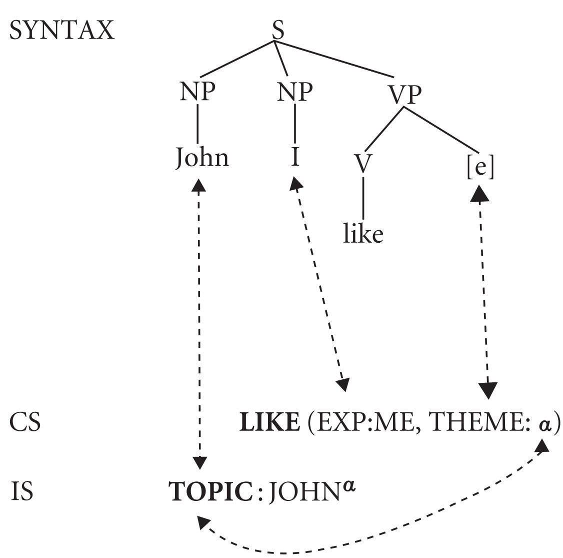


Figure 30.16. Zero or that-relative clause correspondence

(65) a. John$_i$, I like t_i.

b.

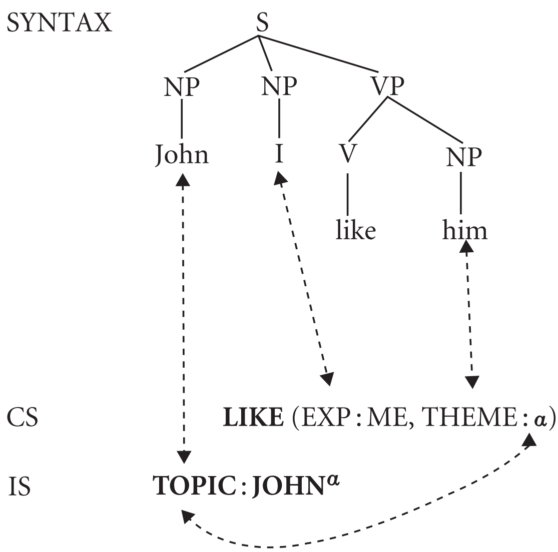


Consider next hanging topicalization, exemplified in (66).

(66) John$_i$, I like him$_i$.

In the correspondence there is a chain whose head is *John* and whose tail is the pronoun *him*.

(67)



The α-argument in CS gets its reference from the topic JOHN$^{\alpha}$ in virtue of being bound by it, as indicated by the superscript α. There is a chain formed between the NPs *John* and *him* that is mediated by the binding of the variable α. Since there is no gap, this chain is not reducible to a movement chain in an account of chains that does not make use of CS.

There are other cases of chains that do not straightforwardly involve gaps and hence lack a plausible analysis in terms of movement. One case is that of implicit arguments. Another involves epithets.

(68) a. (As for) that car$_i$, I really think that the windshield needs to be replaced.
b. (As for) John$_i$, there are very few people that really like that guy$_i$.

The overall picture that emerges is that the key property of a chain is that it links something in syntax to a variable in CS. This variable in turn may correspond to a gap or to an overt expression. The extraction constraints are relevant just in case the chain includes a gap. But the formation of chains is a far more general phenomenon that goes well beyond the set of cases that can be accounted for in terms of movement. The SS step of eliminating A′ movement in the account of chains thus is motivated by the fact that movement is not only unnecessary but insufficient.

30.5 Summary

The Simpler Syntax Hypothesis calls for a radically minimalist theory of syntax. On this approach, only those structures, elements, mechanisms, and constraints that cannot be accounted for in terms of the correspondences between form and meaning are assumed to be components of syntactic theory, formulated in terms of the Parallel Architecture of Jackendoff 2002. I have reviewed a number of applications of this logic and argued that in each case, the SS approach is not only workable but it affords greater empirical coverage than the mainstream analysis.

SS is part of the broader enterprise of understanding how language works, and why it takes the form that it has. A virtue of SS is that it reduces to a significant extent the degrees of freedom available for syntactic description and explanation. As with other "monostratal" accounts, there is only one syntactic representation, and it is considerably simpler than the options that are afforded by accounts that assume invisible functional heads, multiple-level projections, weak and strong features, various types of licensing (before and after "Spell Out") and the like. In the absence of strong positive evidence for constituent structure, a phrase headed by the category V, for example, is flat, regardless of the number of arguments and adjuncts it contains.

In this sense the syntax is radically simpler and the theory is more restrictive, although of course it does place demands on the correspondence rules to get the semantic interpretations right. Our hypothesis is that, in the final analysis, these correspondence rules will prove to be no more complex than the rules needed to interpret richer syntactic structures. To the extent that the cases that I review here support the SS approach, the enterprise appears to be a viable one and one worth pursuing further as an account of how syntax articulates with the other components of a grammatical theory.

This page intentionally left blank

CHAPTER 31

SYSTEMIC FUNCTIONAL GRAMMAR AND THE STUDY OF MEANING

ALICE CAFFAREL

31.1 Introduction

Systemic Functional Linguistics (SFL) is a theory of language that is strongly oriented to the description of how language makes meaning in context (Halliday 1978, 1994, 1996; Halliday and Hasan 1985; Halliday and Matthiessen 1999, 2004). SFL interprets language as meaning potential where all strata of the linguistic system contribute to the making of meaning: the semantic system semanticizes contextual meaning by providing resources to enact and construe it as linguistic meaning; the lexicogrammatical system grammaticalizes this meaning by providing resources

I am greatly indebted to Christian Matthiessen for what I know about Systemic Functional Linguistics in particular and linguistic theories in general. I would also like to thank Kathryn Tuckwell for helping me formulate some of the explanations of theoretical concepts to a non-specialized readership. I am also indebted to Guenter Plum for his editing and comments. Many thanks also to Margaret Hennessy for her suggestions and close reading of the chapter.

to create meaning in wording; and the phonological system realizes meaning by sounding the wordings that realize meaning. This functional orientation—i.e., the orientation to meaning—means that the grammatical analysis of texts in SF terms is not simply a formalized description of the syntax of individual sentences divorced from their co-text (the surrounding language) and context (the relevant extralinguistic activity) but a description of how particular grammatical units are functioning (i.e., making meaning) within particular clauses, within a particular text, and within a particular socio-cultural situation. Thus from an SFL perspective, the study of grammar cannot be carried out independently of the study of meaning; and the interpretation of the meanings construed by the grammar in a particular text is itself informed by the situation and culture in which these meanings were produced, as semantics is "the interface" between grammar and context.

This chapter aims to illustrate the power of SFL and in particular, Systemic Functional (SF) Grammar, as a tool for exploring meaning in any language, as demonstrated in Caffarel et al. (2004). Here we will essentially draw from text instances from French with some references to English. First, we will locate SFL within the wider linguistic community; then we will outline some of the main dimensions and characteristics of SFL and SF grammar; and, to finish, we will explore further aspects of SF grammar in relation to French, discussing what counts as evidence in SFL such as the criteria and "reactances" (Whorf 1956) used for the interpretation of specific categories, and then illustrate the grammar and "crypto-grammar" (Whorf 1956) at work as a meaning-making resource.

31.2 Locating SFL within the wider linguistic community

Halliday shows that, in western thinking about language, there are two main conceptions of language, language as resource and language as rule:

> The sophists saw language as a resource; as a mode of action and a means of putting things across to others. They were concerned with meaning, but not with truth value; if language had any relation to truth this lay in its ability to demonstrate that truth was relative to the believer, and that there was another side to whatever question was at issue. Aristotle saw language as a set of rules; as a mode of judgment and a means of affirming and denying, without reference to other people. He was concerned with truth value, and hence subordinated language to the structure of logic.
>
> So in the earliest flourish of western linguistics we can trace the source of our original metaphor.... We can follow these two strands throughout the subsequent history of

ideas about language in the west. . . . We can identify, broadly, two images of language: a philosophical-logical view, and a descriptive-ethnographic view. In the former, linguistics is part of philosophy, and grammar is part of logic; in the latter, linguistics is part of anthropology, and grammar is part of culture. The former stresses analogy; is prescriptive, or normative, in orientation; and is concerned with meaning in relation to truth. The latter stresses anomaly; is descriptive in orientation; and concerned with meaning in relation to rhetorical function. The former sees language as thought, the latter sees language as action. The former represents language as rules; it stresses the formal analysis of sentences, and uses for purposes of idealization (for deciding what falls within or outside its scope) the criterion of grammatically (what is, or is not, according to rule). The latter represents language as choices, or as resource; it stresses the semantic interpretation of discourse, and uses for idealization purposes the criterion of acceptability or usage (what occurs or could be envisaged to occur). (Halliday 1977: 36–7)

In modern linguistics, we can see the two images manifested quite clearly in two different orientations that have characterized schools of linguistics since the 1920s. The image of language as rule is manifested in formal linguistics; the image of language as resource is manifested in functional linguistics. A partial picture of the two orientations showing the emergence of formal and functional linguistic theories in the twentieth century in rough chronological order is presented in Table 31.1 below.

The differences in orientation between formal and functional theories of language can be summarized by what Halliday (1978) calls the “intra-organism” and

Table 31.1. A partial picture of formal and functional theories of language

	language as rule formal linguistics		language as resource functional linguistics		
1920s–1930s				Prague School	
1940s–1950s	Bloomfieldians => Neo-Bloomfieldians			=> Martinet’s functionalism	British contextualism (Firth and followers)
late 1950s–1960s	Chomsky’s formal grammar: transformational grammar (Syntactic Structures, Standard Theory [Aspects Model])				**Halliday’s Systemic Functional Linguistics**
1970s	(Revised) Extended Standard Theory				
late 1970s–1990s	Government and Binding	Reaction against Chomsky’s particular version: GPSG, LFG, HPSG	Reaction against Chomsky’s formalism: West-Coast functionalism	=> French functionalism: Claude Hagège, Pottier => Dutch functionalism: Simon Dik	

"inter-organism" perspectives. The latter represents the functional view of language and means studying language as social interaction, language as "human behavior"; the former represents the psychological view of language put forward by syntactic theories—it means studying language as knowledge:

> The study of language as knowledge is an attempt to find out what goes on inside the individual's head. The questions being asked are, what are the mechanisms of the brain that are involved in speaking and understanding, and what must the structure of the brain be like in order for the individual to be able to speak and understand language, and to be able to learn to do so. (Halliday 1978: 13)

In the midst of these opposite but potentially complementary approaches to language, more "functionally" oriented grammars evolved from mainstream formal linguistics. In the late 1960s, a framework known as generative semantics arose as an antidote to the Standard Theory. In the late 1970s, formal theories of grammar that were more semantically oriented (in particular toward ideational meaning, to put it in terms of the metafunctions of SF theory discussed in section 31.3.3), such as Generalized Phrase Structure Grammar (GPSG; Gazdar et al. 1985) and Lexical-Functional Grammar (LFG; Kaplan and Bresnan 1982), were developed. This trend was continued with Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994), developed out of GPSG and other work, such as knowledge representation in the 1980s.

Functional interpretations of language were pioneered by Vilem Mathesius and other Prague School linguists before World War II; they were concerned with the functions of language in a general sense, including intellectual and aesthetic functions—what Martin (1990) has identified as extrinsic functionality—as well as functions in the particular sense of principles underpinning the organization of language—what Martin has identified as intrinsic functionality. The major contribution to intrinsic functionality was the focus on the "textual" function (to use Halliday's term, see section 31.3.3 below)—what came to be known as functional sentence perspective. Prague School work was continued after the war, with further contributions on the textual function, with proposals regarding the functions and levels of language, and so on. While the French linguist André Martinet had been influenced by the Prague School in the 1930s, a new generation of European functionalists influenced by the Prague School emerged later, including Simon Dik's (1978, and many other publications) Dutch school of Functional Grammar, and French functionalists such as Claude Hagège (1985) and Pottier (1992). At the same time, the Prague School also influenced Halliday's work on SF theory. However, SF theory differed from continental functionalism in Europe in that its immediate origin was the kind of meaning-based contextualism and system-structure theory developed by J. R. Firth (see 1968) in the 1930s–1950s in Britain.

Functionalism on the other side of the Atlantic has many features in common with European functionalism, but the context in which it developed from the 1970s onward was quite different from the European one. While European functionalists could build on earlier Prague School work or British "contextualism", American functionalists had to establish themselves in reaction against the Chomskyan paradigm. For example, this is how one of the leading American functionalists, Givón (1982: 6–7; see also this volume), characterizes the scene and his own work in relation to it (footnoted references omitted):

The restrictive dogma that came out of the Bloomfieldian amalgam lasted into the mid 1950's. Occasional islands of common sense in the **Traditional Grammar** vein of Jespersen, such as Sapir's mentalism-cum-typology, the functionalism of the **Prague School** and Bolinger, Greenberg's cross-language typological universalism with a diachronic twist, or Pike's cross-disciplinary version of relevance, remained isolated in the calm ocean of post-Bloomfieldian smugness. Linguistics was on its way to exhaustively describe its subject matter, so it seemed.

The **transformational-generative** revolution of the late 1950's represented a vigorous shaking up of the linguistic crucible, yielding a blend of disparate philosophical and methodological features culled from many acknowledged and unacknowledged antecedents. Whether the particular mix and its coherence or lack thereof were the product of design or accident is still a matter of debate.

While functionalists within the systemic tradition have given priority to developing comprehensive descriptions of particular languages (Caffarel et al., 2004), West-Coast functionalists have also given a high priority to typological work (drawing, as Givón notes, on the Greenbergian approach) and to functional accounts of universals (in response to Chomsky's preoccupation). Their studies have increasingly drawn on discourse, providing discourse-based accounts even of the systemic differentiation of word classes, as in Hopper and Thompson's influential work on nouns and verbs:

The properties that prototypical nouns and verbs possess might be thought to be semantic, along the lines of the discussion above. A prototypical verb would then perhaps be one which denoted a concrete, kinetic, visible, effective action, carried out by and involving participants. A prototypical noun might be considered to be one which denoted a visible (tangible, etc.) object. These semantic parameters would correspond to Brown's observation about the primacy of such categories in child language . . . , and to Rosch's assumption that prototypical instances of categories are acquired earlier. Furthermore, the easiest grammatical assignments to make of words to lexical categories are cross-linguistically of these kinds. Concrete, stable things are invariably assigned to the class of Noun, and kinetic effective actions to the class of Verb.

However, although such semantic features are needed, they do not appear to be sufficient for assigning a given form to its lexical class. Prototypicality in linguistic categories depends not only on independently verifiable semantic properties, but also, and perhaps more crucially, on linguistic function in discourse. . . . In this paper we hope to present evidence that the lexical semantic facts about nouns and verbs are secondary to their *discourse roles*. (Hopper and Thompson 1985: 155–6)

Although SFL shares ideas with European functionalists inherited from the Prague School, such as a view of language as a tristratal system, as well as domains of inquiry with West-Coast functionalism, such as typology and discourse studies, it has many distinguishing aspects that spring from Firth's influence.

Firth's ideas about language are central to the development of SF theory as a theory of language as meaning potential in context, which sets it apart from other theories. Recognizing that language as a whole is a resource for meaning and prioritizing a systemic organization over a structural one is what makes SF theory unique among both formal and functional theories. Firth's contextualized theory of meaning and notion of system provided the core ideas for the development of SFL and register theory within SFL, "a theory of functional variation of the general system correlated with contextual variation" (Matthiessen, 1993: 223), which theorizes the connection between language and context.

Here we will take a journey through the metalinguistic potential of SF theory for exploring meaning in context in a particular language with grammar as starting point.

31.3 Main semiotic dimensions and concepts of SFL

SFL provides a very elaborate multidimensional semiotic space (see Halliday and Matthiessen 2004: 21) within which the analyst can explore and describe the creation of meaning in context in a particular language.

31.3.1 Stratification and instantiation

Stratification and instantiation are key dimensions in SFL and provide different perspectives from which to explore language (see Figure 31.1). **Stratification** refers to the organization of the linguistic system in context into strata (context, semantics, lexicogrammar, and phonology/graphology). The relationship between the strata is one of **realization**. Matthiessen (1993: 251) points out that this realizational relationship is dialectic: "... in realizing context, language construes it and in realizing semantics, grammar construes it". Both semantics and lexicogrammar form the content strata of language, and it is this stratification of content in human language that makes it so powerful as a semiotic system, allowing us to expand our meaning potential and linguistic activities:

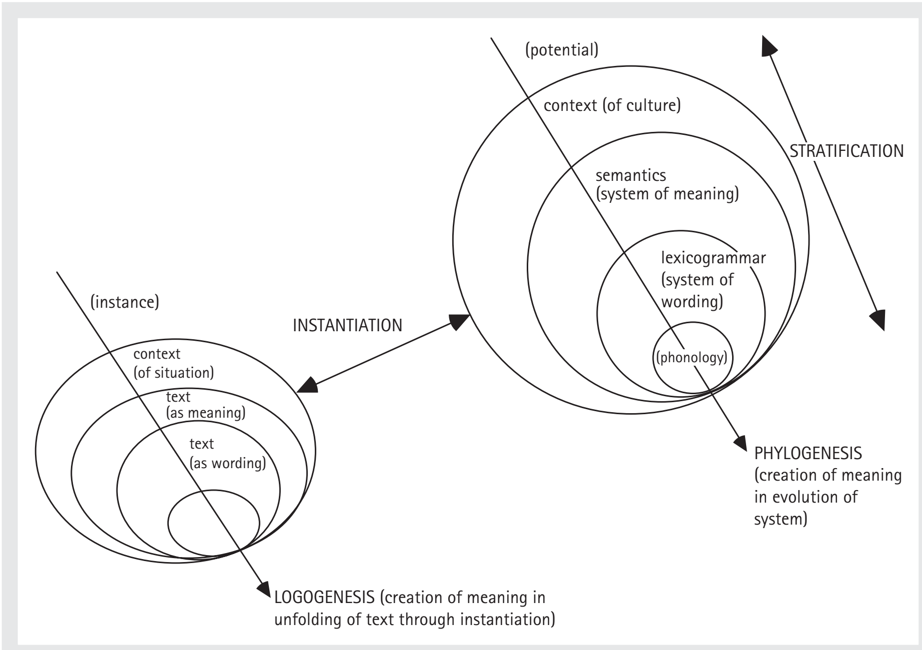


Figure 31.1. Stratification and instantiation (from Caffarel 2006*a*: 9)

We use language to make sense of our experience, and to carry out our interactions with other people. This means that the grammar has to interface with what goes on outside language: with the happenings and conditions of the world, and with the social processes we engage in. But at the same time it has to organize the construal of experience, and the enactment of social processes, so that they can be transformed into wording. The way it does this is by splitting the task into two. In step one, the interfacing part, experience and interpersonal relationships are transformed into meaning; this is the stratum of semantics. In step two, the meaning is further transformed into wording; this is the stratum of lexicogrammar. This is, of course, expressing it from the point of view of the speaker, or writer; for a listener, or reader, the steps are the other way around.

(Halliday and Matthiessen 2004: 24–25)

Instantiation is a cline that provides two perspectives from which to look at language, from the potential (the system) or from the text (the instance):

What we call the "system" is language seen from a distance, as semiotic potential, while what we call "text" is language seen from close up, as instances derived from the potential. In other words, there is only one phenomenon here, not two; langue and parole are simply different observational positions. (Halliday 2005: 248)

Both the relationship between strata and between the system and the instance are interpreted as dynamic. SFL theorizes the relationship between grammar and meaning as natural: lexicogrammatical patterns do not simply "express" pre-existing meanings, rather, grammar is viewed as a resource for making meaning, and meaning is both construed and constructed by the lexicogrammatical patterns of a text. This natural realizational relationship between meaning and grammar will be illustrated in section 31.4.

SFL theorizes the relationship between the grammatical system of a particular language (such as English or French) and individual texts produced by that system as one of "instantiation"—again, a dynamic relationship—in which each instance of language use contributes to our notion of the overall "system", and it is in reference to this that particular lexicogrammatical and semantic selections evidenced in the text have value. Between the system and the instance we find an intermediate region, that of register, which evolved from Firth's contextualized theory of meaning and notions of "restricted languages" (Firth 1968: 124). Matthiessen (1993: 223) views register as the synthesis of a monosystemic and a polysystemic approach:

> To idealize the picture, we can interpret the development of current register theory as a dialectic sequence.
>
> The **thesis** is that language is monosystemic—[...] it seems to be the default in mainstream work. [...] The **antithesis** is Firthian polysystemicness, with restricted languages as the seed for register systemic theory. The uniformity of a single global system is replaced by the diversity of a plurality of more local systems. The **synthesis** is register theory in systemic linguistics—a theory of functional variation of the general system correlated with contextual variation. (Matthiessen 1993: 223)

Depending on the end of the cline of instantiation from which we approach register, instance, or potential, we can interpret it either as a cluster of similar texts, a text type, or as a subpotential of the overall potential specific to a particular situation type, as shown in Figure 31.2.

> Instantiation is a cline, with (like lexicogrammar) a complementarity of perspective. I have often drawn an analogy with the climate and the weather: when people ask, as they do, about global warming, is this a blip in the climate, or is it a long-term weather pattern?, what they are asking is: from which standpoint should I observe it: the system end or the instance end? We see the same problem arising if we raise the question of functional variation in the grammar: is this a cluster of similar instances (a "text type", like a pattern of semiotic weather), or is it a special alignment of the system (a "register", like localized semiotic climate)? The observer can focus at different points along the cline; and, whatever is under focus, the observation can be from either of the two points of vantage. (Halliday 2002: 412)

Thus, with regard to instantiation, it is also possible to describe varieties of language or registers that are intermediate between the overall system and the individual instance by describing a number of individual instances that are produced in similar contexts (to use the weather/climate analogy again, by studying the weather in a number of desert regions, we can generalize about desert climates as a sub-system of the "global climate"). Such varieties are not qualitatively different from each other,

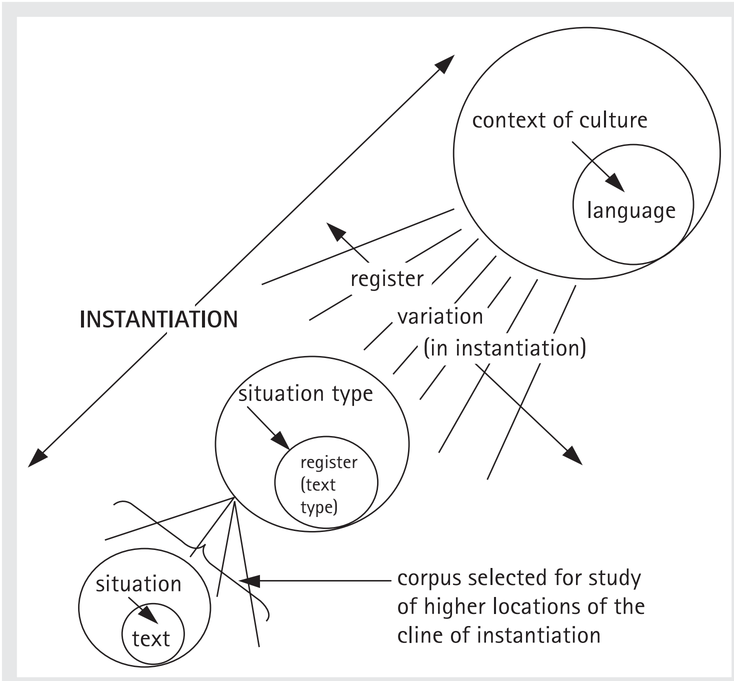


Figure 31.2. Register variation and instantiation (from Caffarel 2006*a*: 12)

except in that they are quantitatively (probabilistically) different—i.e., there is no categorical difference in the linguistic options available in different contexts, but the probability that particular options will be chosen varies across text types. To take a simple example, in English there are three basic choices (categories) of tense (past, present, future) and, in general, looking across a number of text types, future tense is used about ten percent of the time; but in the sub-system of English that we could call "the register of weather forecasting", future tense is used in about ninety percent of clauses, and this is one of the probabilities that sets it apart as a register (Halliday 1996). The SFL theory of register thus provides a useful framework for investigating the similarity (or otherwise) of lexicogrammatical choices in texts.

31.3.2 Logogenesis, phylogenesis, ontogenesis

While logogenesis is the synchronic creation of meaning in a text through instantiation by means of choosing particular features from the overall system, phylogenesis refers to the creation of meaning seen from a diachronic systemic perspective, looking at the evolution of the system from an historical perspective. The construction of meaning can be seen as a continuous semiotic act which began in an undetermined far remote past, is in progress at present, and will continue in the future (phylogenesis). And, within this progressive semiotic act that represents the evolution of language as meaning potential through time, take place an infinite number of semiotic acts that represent the series of linguistic texts that are instantiated at each moment by the "speaking" world (logogenesis). Another type of genesis in the creation of meaning, of importance here, is ontogenesis or the evolution of the linguistic system in the individual from protolanguage to language. As Matthiessen (1995: 48) points out:

> Both ontogenesis and phylogenesis give us perspectives on the development of lexicogrammar as a meaning-making resource. In the ontogenetic perspective, we can actually study how lexicogrammar emerges as a stratum intermediate between semantics and phonology. It begins to emerge as the child moves from his/her proto-language into the mother tongue: see, e.g., Halliday (1975) and Painter (1984). This perspective helps us understand not only the motivation for the development of lexicogrammar but also the origin of its metafunctional organization. In the phylogenetic perspective, we can only track a very short period of the history of language in the human species—far too short to tell us anything about the emergence of lexicogrammar. We can, however, explore how particular systems have changed and how whole groups of systems have changed in resonance with one another. Such evolutionary change is typically gradual; and we can interpret it as gradual change of systemic probabilities. (Matthiessen 1995: 48)

The most important process in the genesis of language from the protolinguistic to the linguistic stage is the shift from a bistratal linguistic system (content: expression) to a tristratal linguistic system (semantics: lexicogrammar: phonology), whereby content is stratified into semantics and grammar.

The ontogenetic development of grammar increases our meaning potential by allowing for metafunctional diversification, i.e., for producing multiple meanings simultaneously as a single structural frame. In addition, the stratification of content gives the means to depart from the congruent to arrive at the incongruent/metaphorical (meaning through grammatical metaphor); see section 31.4.

31.3.3 The metafunctional dimension

One key aspect of the organization of the content levels of language is **metafunctional diversification**, which allows the simultaneous creation of multiple meanings: ideational, which comprises both experiential and logical meanings (creation of text/clause as experience (experiential meaning) and sequences of experiences (logical meaning)), interpersonal (creation of text/clause as an exchange), and textual (creation of text/clause as a message in context). Choices within the three metafunctions (ideational, interpersonal, and textual) are to some extent dependent on the tri-dimensional structure of context formed by the Field (the activity taking place), Tenor (role and status relationships among participants) and Mode (rhetorical function of language/type of language). What we call a register within SFL (see section 31.3.1) can be seen as the semantic correlate of a configuration of these three contextual variables, and from the contextual information provided by these variables we can make predictions about the linguistic features of a particular text type.

Hasan (1989: 237) points out that:

> Halliday does not claim a simple and absolute correspondence between *some specific* metafunction and *some specific* contextual variable, as if one mirrored the other; he does, however, claim that typically the ideational metafunction is constitutive of field, the interpersonal of tenor and the textual of mode. (Hasan 1989: 237)

Caffarel (2006*b*: 209–11) illustrates the correlation between contextual variables and linguistic choices within the three metafunctions in the context of a French linguistics lecture by showing:

> ... how the "institutional setting", i.e. the teaching of linguistics at university, influences choices in transitivity and lexis (ideational meaning), the "relationship between participants", i.e. the variation in status and knowledge between students and teacher, influences the dynamics of the exchange in the classroom and the choices in mood (interpersonal meaning), "the channel of communication adopted", here spoken instruction, influences textual choices and the organization of the flow of information.
>
> The following is an example of the kind of text the teacher might offer in the early part of the semester:
>
> Text 1:
>
> *Un aspect important de la théorie SF est la relation naturelle qui existe entre sémantique et lexicogrammaire. Les choix lexicogrammaticaux servent à réaliser des choix sémantiques. La relation qui existe entre les deux est parfois congruente, parfois métaphorique. C'est le fait que grammaire et sémantique ne sont pas toujours alignés qui rend la langue un outil de communication flexible et puissant.*
>
> *Toute énonciation est en partie prévisible à partir du contexte, du Champ du discours (l'activité dont on parle), de la Teneur du discours (les relations qui existent entre interlocuteurs) et du Mode du discours (le rôle du langage/type de langage).*
>
> [An important aspect of SF theory is the natural relationship that exists between semantics and lexicogrammar. Lexicogrammatical choices serve to realize semantic choices. The relationship that exists between the two is sometimes congruent and sometimes metaphorical. It is the fact that grammar and semantics are not always aligned that makes language a flexible and powerful communication tool.
>
> Each text instance is in part predictable from context, from the Field of discourse (the activity taking place), the Tenor of discourse (the nature of the relationship between participants) and the Mode of discourse (the role/type of language).] (my translation).
>
> (Caffarel 2006*b*: 209–11)

The kind of linguistic choices that can be predicted from a "lecture-type" classroom situation and are found in Text 1 are summarized in the table below:

Table 31.2. Correlation between contextual variables and metafunctions (adapted from Caffarel 2006*b*: 211)

Contextual variables	Linguistics features
Field: "What is taking place" Teacher introducing students to theory and ideas about language	Ideational choices: Participants: linguistic concepts Transitivity: relational clauses: defining, naming, describing
Tenor: "Who is taking part" The lecturer, "linguistic expert" is giving information to students, "non-expert in linguistics".	Interpersonal choices: Mood: declarative. Speech roles: Lecturer: speaker/giver of information Students: listener/demander of information
Mode: "What role language is playing" The function of the text is to present and explain ideas about language in a French linguistics class.	Textual choices: The language of instruction is French, the medium is spoken, and the language "written", with a high lexical density and no clause complexes. The themes are mostly unmarked and the text tends to have a linear progression where the Rheme of the previous sentence becomes the Theme of the following sentence. This kind of thematic progression is typical of explanatory texts.

Thus the metafunctional dimension of the systemic organization of language enables the analysis of the construal of context by language—how texts as meaning and wording realize particular contextual configurations of Field, Tenor, and Mode—and the consequent identification of sub-potentials of the overall potential as schematized in Figure 31.3.

Sub-potentials of the overall potential, or register-specific systems, can be identified by analyzing clusters of instances that make similar choices from the three metafunctional components, choices that are themselves related to the Field, Tenor, and Mode of discourse. We will return to the metafunctions in section 31.4, where we explore the grammatical stratum in more detail. To finish this overview of key SF notions, we will turn to the systemic dimension.

31.3.4 The systemic dimension: Paradigmatic organization

From the discussion so far it transpires that the notion of system is central to SFL. Indeed, since language is a meaning potential, it is the system that provides the semiotic environment for structural description in SF theory. Each stratum of the linguistic system is viewed as a network of interrelated options or alternative choices

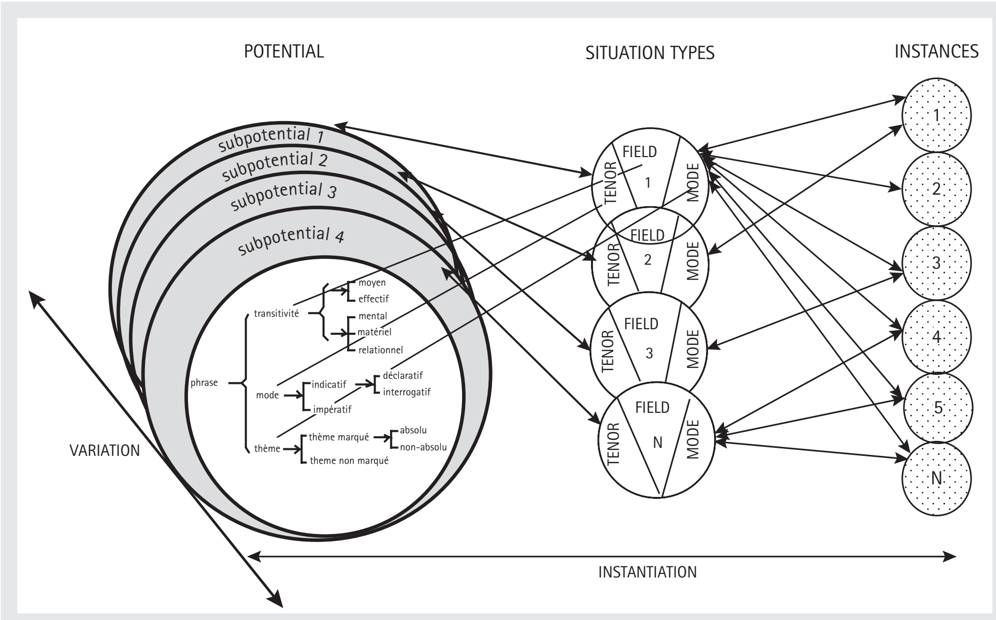


Figure 31.3. Variation in meaning and contextual variation

Note: English labels for the features are: phrase = clause; transitivité = transitivity; moyen = middle; effectif = effective; mental = mental; matériel = material; relationnel = relational; mode = mood; indicatif = indicative; déclaratif = declarative; interrogatif = interrogative; thème = theme; théme marqué = marked theme; thème non-marqué = unmarked theme; absolu = absolute; non-absolu = non-absolute.

called systems, so the particular choices that are taken up have value relative to the choices not taken up. The systems of options in each stratum can be drawn up as "system networks" (see Figure 31.4) which allow analysts to make systematic and statistical statements about the composition of a text (or corpus of texts), and are also particularly adaptable to the computational modeling of text. The systems within the networks are ordered according to the scale of delicacy from most general to most delicate systems, with lexis being interpreted as most delicate grammar (Cross 1993; Halliday 1961, 1966; Hasan 1987), hence the term "lexicogrammar" rather than simply "grammar":

> The lexical system is not something that is fitted in afterwards to a set of slots defined by the grammar. The lexicon . . . is simply the most delicate grammar. In other words, there is only one network of lexicogrammatical options. And as these become more and more specific, they tend more and more to be realized by the choice of a lexical item rather than the choice of a grammatical structure. But it is all part of a single grammatical system. (Halliday 1978: 43)

The system represented in Figure 31.4 is a fragment of the system of MOOD for the French clause. It reads as follows: when entering the French MOOD system, one may choose the option "indicative" or the option "imperative". The feature "indicative" becomes the entry condition for a more delicate system offering the

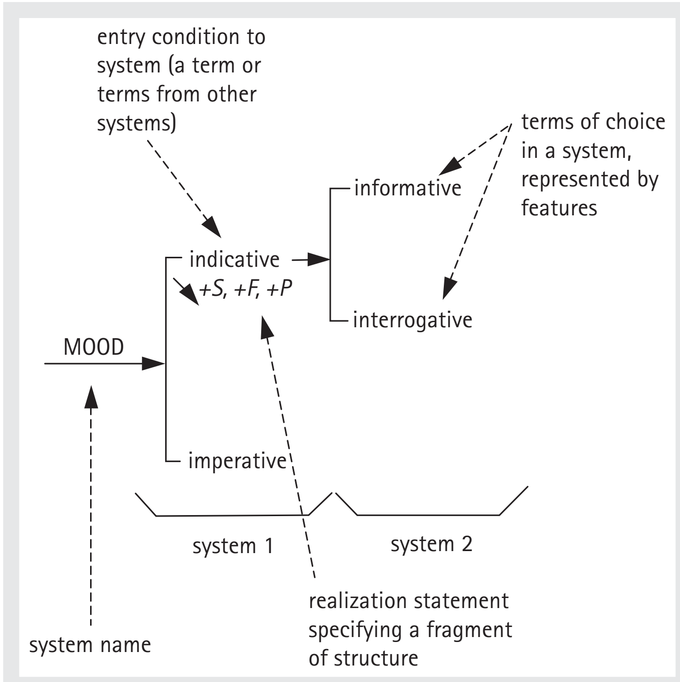


Figure 31.4. The representational conventions of a system network (from Caffarel 2006*a*: 13)

choice between "informative" or "interrogative". The realization statement under the feature "indicative" indicates that for the realization of an indicative clause in French there must be a Subject, a Finite, and a Predicator. While square brackets mean "or", curly brackets mean "and", thus enabling the representation of simultaneous systems as shown in Figure 31.5 in the next section.

This section introduced some of the main dimensions in SFL from which to approach, describe, and understand language and meaning. We will now focus on the lexicogrammatical stratum in more detail and on what makes it such a powerful tool for the exploration of meaning.

31.4 Systemic Functional Grammar

As discussed in section 31.3.3, one crucial aspect of the organization of language and more precisely of the content level of language (semantics and lexicogrammar) is metafunctional diversification. A clause means simultaneously in three different

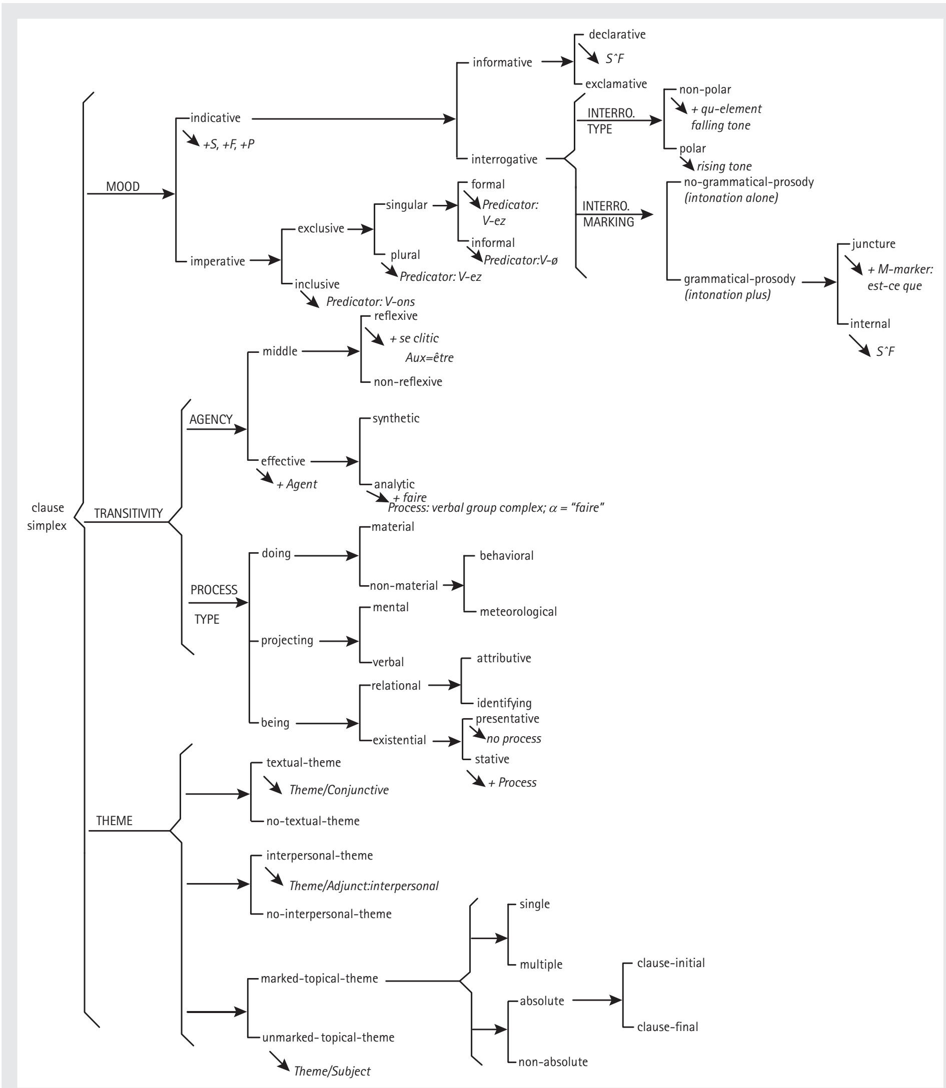


Figure 31.5. A simplified system network of the French clause

ways: it functions as part of an exchange of information or goods and services (interpersonal metafunction), it represents a reality or series of realities (ideational (logical and experiential) metafunction), and it organizes these meanings cohesively in relation to the text and context in which the clause is produced (textual metafunction). Thus if we ask, what does *She is coming back tomorrow* mean, we can answer by explaining that it is a statement. Its function is to give some information to the listener. It is the response to a question like, *When is she coming back?* Also, the clause is about an event. It is about a participant (*she*) involved in a process (*to come* *back*) located in time (*tomorrow*). In addition to being an exchange, establishing and maintaining social relations between speaker and listener and representing an element of experience (an event), the clause is presenting information that coheres with a particular co-text and context. Here the use of the personal pronoun *she* tells us that its referent can be retrieved from the co-text or context or both, and that the speaker and the listener both have shared knowledge of that person. What is important textually is typically located at the beginning of the clause in English. It is what we call the Theme. Here, *she* is the Theme of the clause, the point of departure of the message, but it does not need to be. *Tomorrow* could have been selected as Theme, as in *It's tomorrow that she is coming back* (marked predicated Theme). As a structure above the clause, we recognize the clause complex whereby sequences of events are linked logically, as in *She came back and left again* in which two experiential configurations are paratactically related through a logico-semantic relation of extension. But grammar is a system that makes meaning not only through metafunctional diversification (multiple modes of meaning) but also through grammatical metaphor (which increases the power of grammar as a meaning-making resource) and through foregrounding and de-automatization whereby latent/covert patterns and interrelations of patterns make meanings of a second order, free from the control of semantics. All these aspects of grammar as a meaning-making resource will be illustrated in sections 31.4.1.1, 31.4.1.2, and 31.4.1.3 in relation to MOOD, TRANSITIVITY, and THEME.

31.4.1 A metafunctional slice of French grammar

Grammatical systems that contribute to construing the different types of meanings—logical, experiential, interpersonal, and textual—are the CLAUSE COMPLEX systems (ideational: logical metafunction), the TRANSITIVITY systems (ideational: experiential metafunction), the MOOD, MODALITY, and POLARITY systems (interpersonal metafunction), and the THEME and VOICE systems (textual systems). Here we will only discuss some features of TRANSITIVITY, MOOD, and THEME systems as instantiated in the French language. (For a discussion of clause complex systems and other systems in French, see Caffarel 2006*a*; for a discussion of English systems, see Matthiessen 1995, and Halliday and Matthiessen 2004; and for a typological overview of clausal systems across languages, see Caffarel et al. 2004.)

From a syntagmatic perspective, features from the MOOD, TRANSITIVITY, and THEME systems are realized in three simultaneous layers of functional structuring corresponding to the three metafunctions. A description of an example thus includes both a systemic aspect (the features selected, or "selection expression") and a structural aspect (the structural functional specifications realizing the systemic features).

Table 31.3. Metafunctional analysis of (1)

	La fenêtre	s'est	ouverte
Systemic features of clause:	{indicative: declarative:...; material & middle:...; unmarked theme & active...}		
Interpersonal: modal structure	Subject	Finite	Predicator
Ideational: experiential: transitivity structure	Actor/Medium	Process: material	
Textual: thematic structure	Theme	Rheme	

A metafunctional analysis of clauses (1) and (2) follows:

(1) La fenêtre s'est ouverte
The window opened

(2) Louis a ouvert la fenêtre
Louis opened the window

The boxed representation of the multilayered clause structures shows that these two clauses vary primarily in their choices of transitivity features. While in (1) the constituent that functions as Subject conflates with the Actor/Medium, in (2) it conflates with the Actor/Agent. This difference foregrounds that in (1) the event is

Table 31.4. Metafunctional analysis of (2)

	Louis	a	ouvert	la fenêtre
Systemic features of clause:	{indicative: declarative:...; material & effective:...; unmarked theme & active:...}			
Interpersonal: modal structure	Subject	Finite	Predicator	Complement
Ideational: experiential: transitivity structure	Actor/Agent	Process: material		Goal/Medium
Textual: thematic structure	Theme	Rheme		

represented as self-engendered, while in (2) it is represented as brought about by Louis. In addition, while in (1) it is the Medium that functions as Theme, i.e., as point of departure of the message, in (2) it is the Agent. In sections 31.4.1.2 and 31.4.1.3, where we analyze and interpret the meanings of a text, we will see that the properties of the clause can be linked to its contribution to the unfolding discourse, in terms of thematic progression, for example, and to its contribution to higher order meanings, social or ideological, for example.

This distinction between clauses with an external Agent as in (2) and clauses with no external Agent as in (1) reflects the distinction between [effective] and [middle] in the TRANSITIVITY system (see Figure 31.5).

The clause is thus the point of departure for any grammatical analysis of a text. In order to analyze the metafunctional mappings of a text, the text will need to be divided into clauses [clause simplexes as well as clauses within clause complexes], the clause being the highest-ranking unit in the grammar above the group/phrase-word-morpheme ranks:

> According to the rank scale (the hierarchy of units), a unit of a particular rank will serve to realize a functional element in a unit of the rank immediately above; a group will serve to realize an element of clause structure, a word will serve to realize an element of group structure, and a morpheme will serve to realize an element of word structure. However, a unit may come to serve to realize an element of a unit of the same rank or of lower rank. This is called **rankshift**. (Matthiessen 1995: 21)

For example, in the following clause, *I like people [[who are tolerant]]*, "who are tolerant" is a rankshifted or downranked clause in that it functions at group rank rather than clause rank and serves to modify the nominal group rather than expand the clause through interdependency and logico-semantic relations.

We will now explore key aspects of a French clause, as exchange, representation, and message, from both a paradigmatic and syntagmatic viewpoint.

31.4.1.1 *French MOOD systems at a glance*

The primary semantic speech functions of "giving and demanding information or goods and services" that are essential to an exchange can be assumed to be

Table 31.5. Key systems of the French clause

Rank	Metafunctions		
	Interpersonal	Experiential	Textual
Clause	as exchange MOOD	as representation TRANSITIVITY	as message THEME

general to all languages. From this assumption, it follows that we can predict that the most general MOOD options that realize the speech functional semantics of command, statement, and question will be similar across languages (see Matthiessen 1995; Caffarel et al. 2004; and Teruya et al. 2007, for an overview of Mood options across languages). Indeed, in Caffarel et al. (2004) we found that the grammar of MOOD in French resembles that of English and other languages in terms of its primary MOOD options for the free clause, differentiating between indicative and imperative as well as between the indicative subtypes, declarative and interrogative. However, as our description of the interpersonal grammar of French became more delicate, the MOOD options in the systems were found to be more specific to French and the structural realizations of these options were also found to differ from English and other languages with French making use of all three types of prosody found across languages: intonational prosody, juncture and internal prosodies (grammatical prosody).

If we look at the MOOD system in Figure 31.5 above and in particular at the interrogative MOOD options, we can see that in French there are three possible ways of marking that a clause is interrogative: with intonation alone, with the Mood marker *est-ce que*, or by inverting the order of Subject and Finite. In order to work out when one structure is used rather than another we would have to look at their uses in context. In fact, the different prosodic modes of marking an interrogative in French are very much register specific and in particular linked to the contextual variable of Tenor. Thus, the feature “no-grammatical-prosody” tends to be selected in a spoken context where interactants are familiar with each other and young; “grammatical-prosody: internal” tends to be used in written contexts and formal spoken contexts where there is distance (in age, social status, etc.) between interactants; while “grammatical-prosody: juncture” tends to be used in both spoken or written mode but in informal contexts where interactants of all ages are close.

Within the SFL model, the description of particular structural properties is typically argued from discourse. Thus the exploration of dialogic texts provides evidence for how the negotiation is carried forward in French or any other language and the Mood options realized.

Consider the following “information” exchange:

(1) Jeanne

Ce bruit,	**tu**	**l’**	**as**	**entendu?**
	Subject	Complement: Clitic	Finite	Predicator
Absolute Theme (see section 31.4.1.3)	**Negotiator**⟶			
This noise	*you*	*it*	*have*	*heard*

‘Did you hear this noise?’

(2) Paul

J'	**ai**	**entendu**	quelque chose de bizarre, comme une explosion.
Subject	Finite	Predicator	Complement
Negotiator→			Remainder
I	*have*	*heard*	*something strange, like an explosion*

'I heard something strange, like an explosion.'

(3) Jeanne se tourne vers Louis et dit: (Jeanne turns towards Louis and says:)

Et	toi,	**tu**	**as**	**entendu?**
Textual Theme	**Absolute Theme**	Subject	Finite	Predicator
		Negotiator→		
and	*you,*	*you*	*have*	*heard*

'And, did **you**?'

(4) Louis

Moi,	Non.
Absolute Theme	Polarity-marker
Me	*no*

'I didn't.'

(5) Jeanne à Paul et Louis (Jeanne to Paul and Louis):

Vous	venez?
Subject	Finite/Predicator
Negotiator→	
you-plural	*come*

'Are you coming?'

(6) Louis

Non,	**je**	**reste.**
Polarity-marker	Subject	Finite/Predicator
	Negotiator→	
No	*I*	*stay*

'No, I'm staying.'

(7) Paul

(a)

Tu	restes,	toi!
Subject	Fin/Pred	Reprise Th.
Negotiator→		
You	*stay*	*you*

'**You** are!

(b)

Nous	on	s'en va.
Abs Th.	Subject	Fin/Pred
		Negotiator→
We	*we*	*go*

We are going.'

This short exchange illustrates some characteristics of the French clause as a move in an exchange. We can see that the resolution of the exchange revolves around the tossing back and forth of Subject, Finite, and Predicator as in (1), (2), and (3) for example, or the ellipsis of these elements as in (4). In contrast, in English only Subject and Finite are essential to the resolution of the exchange as in: "Did you hear it? Yes, I did/No, I didn't." That part of the clause that comprises Subject, Finite, and Predicator is referred to as the Negotiator to distinguish it from the functionally analogous Mood element of English (see Halliday 1994). In French, unlike in English, we can also have clitics as part of the Negotiator that prefix themselves to the Finite as in:

(8) *Est-ce que tu as entendu l'explosion?*

"Did you hear the explosion?"

<table>
<tr><td>Est-ce que</td><td>tu</td><td>as</td><td>entendu</td><td>l'explosion?</td></tr>
<tr><td></td><td>you</td><td>have</td><td>heard</td><td>the explosion?</td></tr>
<tr><td rowspan="3">Mood-
interrogator</td><td>Subject</td><td>Finite</td><td>Predicator</td><td>Complement</td></tr>
<tr><td colspan="3">Negotiator⟶</td><td>Remainder</td></tr>
<tr><td colspan="4">Negotiatory structure⟶</td></tr>
<tr><td colspan="5">Interpersonal structure⟶</td></tr>
</table>

(9) *Oui je l'ai entendue.*

"Yes, I heard it."

<table>
<tr><td>Oui</td><td>je</td><td>l'</td><td>ai</td><td>entendue</td></tr>
<tr><td>Yes</td><td>I</td><td>it</td><td>have</td><td>heard</td></tr>
<tr><td></td><td>Subject</td><td>C-clitic</td><td>Finite</td><td>Predicator</td></tr>
<tr><td></td><td colspan="4">Negotiator</td></tr>
<tr><td>Polarity Adjunct</td><td colspan="4">Negotiatory structure⟶</td></tr>
<tr><td colspan="5">Interpersonal structure⟶</td></tr>
</table>

As (9) shows, once the Complement is Given and pronominalized, it becomes part of the Negotiator, and the Predicator agrees in number and gender with the Complement clitic, thus creating a prosody across the Negotiator. In contrast, as shown in (8), when the Complement is part of the Remainder, there is no agreement between Predicator and Complement.

The foregoing examples have illustrated some of the specificities of French Mood systems and structures. It has also illustrated the natural relationship between semantics and grammar, showing, for example, that the speech function of "question" is congruently realized by the interrogative mood in the grammar. However, grammatical realizations of semantic categories are not always congruent, and a given meaning can be realized in different ways in the grammar by means of metaphorical modes of expression.

> Metaphorical modes of expression are a characteristic of all adult discourse. There is a great deal of variation among different registers in the degree and kind of metaphor that is encountered; but none will be found entirely without it. The only examples of discourse without metaphor that we normally meet with are in young children's speech, and in traditional children's rhymes and songs that seem to survive for that reason: that they lack grammatical metaphor. (Halliday 1994: 342)

An example of MOOD metaphor, for instance, would be realizing a command by an interrogative or a declarative rather than an imperative, depending on context. Thus the command "close the window" could be realized as "could you please close the window" or "I would like you to close the window" depending on the contextual variables of Field, Tenor, and Mode.

31.4.1.2 *French TRANSITIVITY systems at a glance*

As shown in Figure 31.5, there are two simultaneous TRANSITIVITY systems, the PROCESS TYPE and the AGENCY system. These two systems realize two different but complementary semantic construals of reality, what has been referred to by Halliday (e.g., Halliday 1994) as the transitive and ergative models of participation in the process. While the transitive model looks at experience from a particularizing perspective, differentiating different domains of experience realized by different process types in which process specific participants are involved (for example, the Actor in a material clause, the Sayer in a verbal clause, the Senser in a mental clause), the ergative model looks at experience from a generalizing perspective where all process types are seen as alike and can either be represented as having agency (explicit or implicit) or no agency. Operating with complementary transitivity systems means that we can interpret different levels of experiential meanings in texts.

Although Halliday's ergative and transitive categories can be applied to the description of French transitivity and many other languages, we cannot assume that they are theoretical categories that can be applied to the description of transitivity of any language. However, evidence from other languages (see Caffarel et al. 2004) supports the hypothesis that all languages have complementary transitivity systems, one particularizing, the other generalizing.

As Figure 31.6 shows, the transitive model is related to the PROCESS TYPE system in the grammar, while the ergative model is related to the AGENCY system. Again, from an SFL perspective, in order to interpret how reality is realized in a particular language, we look at grammatical patterns in texts, and to establish what kind of process types should be recognized in a particular language, we use

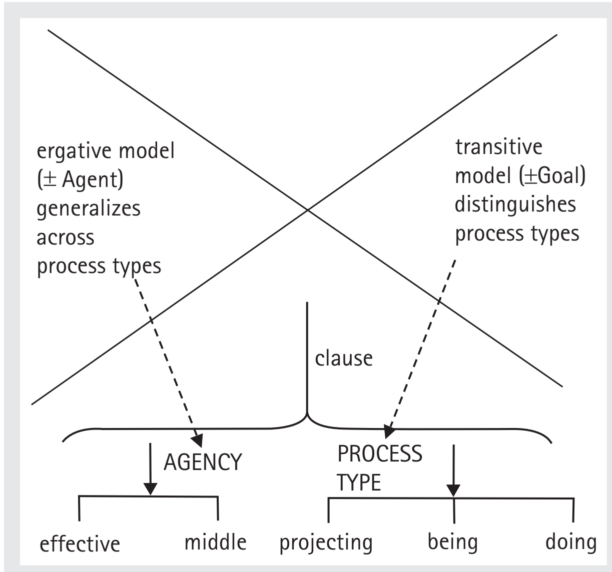


Figure 31.6. Semantic models of participation and transitivity systems (from Caffarel 2006*a*: 66)

criteria or "reactances" in Whorf's use of the term. Thus, as Figure 31.5 shows, in French we recognize three general domains of experience, doing (which includes material, behavioral, and meteorological processes), projecting (mental and verbal processes), and being (relational and existential processes). There are a number of criteria or "reactances" that can be used to differentiate them. These criteria are summarized in Table 31.6 below.

Table 31.6 shows that in French a number of reactances can be used to differentiate the various process types: whether a process can be probed by the general "doing" process "faire"; whether a process can project; whether a process can have a metaphenomenon as participant; whether the Medium has to have consciousness or not; whether the auxiliary "être" or "avoir" is used in the compound tenses; and whether the unmarked past tense for that process is the simple/compound past or the imperfect past.

Consider now the following text, which is analyzed both in terms of the transitive and the ergative perspective, highlighting both process types and process specific participants as well as whether or not a clause has an Agent. The Agent is the participant that causes/brings about the realization of the Process/Medium nucleus. The Medium is the participant essential to the realization of the Process, without which there would not be any process. The Range is the participant that specifies the nature of the process or elaborates the Process. Unlike the Medium in an effective material clause, the Range is neither affected nor created by the Process.

Table 31.6. Criteria for distinguishing process types (adapted from Caffarel 2006*a*: 86)

Grammatical probes	Process type		
	Material (doing)	Mental and verbal (projecting)	Relational (being)
Pro-verb *faire, arriver*	YES *Qu'est-ce qu'il lui a fait? Il l'a frappé*. What did he do to him? He hit him. *Qu'est-ce qui est arrivé? Il a plu*. What happened? It rained.	NO	NO
Project	NO	YES *Elle pense qu'elle réussira*. She thinks that she will succeed.	NO
Meta-phenomenon	NO	YES *Elle pense [[partir en France]]*. She is thinking of going to France.	NO
+ Medium endowed with consciousness	NO *Le train arriva en retard*. The train arrived late.	YES *Elle pensa [[arriver en retard]]*. She thought of arriving late.	NO *Le train était en panne*. The train had broken down.
Auxiliary Unmarked past	Avoir or être Simple/compound past	Avoir Imperfect past	Avoir Imperfect past

The functions of Agent, Medium, and Range are ergative functions and can be applied to all processes.

Text 2

Afghanistan: Deuxième vague de bombardements (LEMONDE.FR|08.10.01 |)[1]

[1] **Afghanistan: second wave of bombings**

(1) A second wave of bombings hit Kabul late on Sunday night.
(2) The first Anglo-American operations started in Afghanistan at 6.15 in the evening (French time).
(3a) George W. Bush confirmed, in an address to the nation on Sunday at 6.50,
(3b) that the United States and Great Britain had commenced "targeted" military strikes.
(4) The Anglo-American armed forces launched a first offensive against Kabul, the Afghan capital, Jalalabad, Kandahar, and Herat.
(5) The spokesperson for the Pakistani Ministry of Foreign Affairs expressed regret that diplomatic efforts had not managed to convince the Taliban leaders to "respond to international demands".

(1) [Agent/Actor] Une deuxième vague de bombardements [Material: effective] a frappé [Medium/Goal] Kaboul, [Circ: time] tard dans la soirée de dimanche.
(2) [Medium/Actor] Les premières opérations anglo-américaines [Material: middle] ont commencé [Circ: place] en Afghanistan [Circ: time] à 18h15 (heure française).
(3a) [Medium/Sayer] George W. Bush [Verbal: middle] a confirmé, [Circ: Place] dans un discours à la nation [Circ: Time] dimanche à 18h50,
(3b) que [Medium/Actor] les États-Unis et la Grande-Bretagne [Material: middle] avaient entamé [Range] des frappes militaires "ciblées".
(4) [Medium/Actor] Les forces armées anglo-américaines [material: middle] ont lancé [Range] une première offensive [Circ: Place] sur Kaboul, la capitale afghane, Jalalabad, Kandahar et Hérat.
(5) [Medium/Senser] Le porte-parole du ministère pakistanais des affaires étrangères [Mental: middle] a regretté [Range: metaphenomenon] [[que les efforts diplomatiques n'aient pu permettre de convaincre les dirigeants talibans de "répondre aux demandes internationales"]].
(6) [Agent/Attributor] Les talibans [relational: attributive: effective] ont qualifié [Medium/Carrier] les bombardements [Range/Attribute] "d'acte terroriste".
(7a) [Medium/Carrier] Ben Laden, [Circ: Place] dans un message préenregistré [[diffusé dimanche soir par Al-Jazira TV]], [relational: attributive: middle] se dit [Range/Attribute] prêt à "la confrontation"
(7b) et [verbal: middle] affirme
(7c) qu' [Medium/Actor] un "groupe de musulmans" [material: middle] a bien commis [Range] les attentats du 11 septembre.

If we look at the transitivity resources of Text 2 on the beginning of the bombing of Kabul in 2001 by the Anglo-American coalition, a number of patterns emerge. From a transitive perspective, we can see that there are a majority of material processes. Three of those mean "to begin" (*commencer*, *entamer*, and *lancer*). The Actor in these three middle material clauses of doing is predictably the Anglo-American coalition. From an ergative perspective, we can see that there is very little usage of agency, and, on the only occasion where the Agent is involved in a material process, this Agent is a nominalization *une deuxième vague de bombardements* ("a second wave of bombings"). Such use of grammatical metaphor, where processes are realized as nominal groups rather than verbal groups, is typical of "hard news" stories where there is typically "an absence of explicit personal and subjective engagement from the journalist" (see Caffarel and Rechniewski 2008).

(6) The Taliban described the bombings as a "terrorist act".
(7a) Bin Laden, in a prerecorded message broadcast on Sunday evening by Al-Jazeera Television, announced himself ready for "confrontation"
(7b) and claimed
(7c) that a "group of Muslims" had indeed carried out the attacks of September 11. (translated by Margaret Hennessy)

Table 31.7. Summary of transitivity resources in Text 2

Agent	Medium	Process	Range
Vague de bombardements	Kaboul	material	
	Opérations anglo-américaines	material	
	Bush	verbal	
	Les États-Unis et la Grande-Bretagne	material	frappes militaires
	Les forces armées anglo-americaines	material	première offensive
	Le porte-parole pakistanais	mental	les efforts diplomatiques
Les talibans	Les bombardements	relational	acte terroriste
	Ben Laden	relational	prêt à la confrontation
	[Ben Laden]	verbal	
	Un groupe de musulmans	material	attentats

The journalist is supposed to remain neutral or not take a position but rather provide factual information. The factuality of this event is emphasized in the text by a number of circumstances of Time and Place. However, the representation of the event itself, "the bombing of civilians by the Anglo-American troops", carries implicitly ideological meaning by construing the Agent as an event impacting on a city, here Kabul, rather than as people killing other people. The Anglo-Americans are on the other hand represented as Actor/Medium; they are acting and doing but they are not represented as acting upon other people or affecting them. This is a subtle way of not assigning responsibility to the "bombers". The use of grammatical metaphor, through nominalizing the action of bombing, has the effect of diminishing the actions of the coalition and distancing the reader from the atrocities of these actions. By not pointing the finger at any particular group, the journalist is not explicitly assigning responsibility to anyone in particular and by doing so gives an appearance of neutrality.

In the next section on THEME systems, we will see that the thematic analysis of the same text converges toward a semblance of neutrality toward the event.

31.4.1.3 *French THEME systems at a glance*

Figure 31.5 showed that the French clause can select Themes from the three metafunctional components, Textual, Interpersonal, and Experiential (Topical), as can the English clause (see Halliday 1994). Textual Themes are discourse markers, conjunctives, and continuatives, as "and" in "and she kept on writing"; Interpersonal Themes are interpersonal adjuncts (modal adjuncts, comment adjuncts, polarity adjuncts) as "probably" in "Probably, she will be late again", as well as vocatives and Mood markers; Topical Themes conflate with a function in the transitivity structure and are either a participant, a circumstance, or a process. While there is not always a Textual or Interpersonal Theme, a Topical Theme is essential to the realization of the clause as a message. It is the local context of the clause, its point of departure. In a declarative clause, it can be the subject (unmarked Topical Theme) or the complement or a circumstance (marked Topical Theme). In addition, French can also have Absolute Themes, i.e., Themes that do not have a role in the experiential or interpersonal structure but only carry textual meaning; they are typically coreferential with a participant in the clause. Thus from a realizational viewpoint, French can be seen to have properties of head marking languages, where the process is marked for case and participants are introduced at the beginning of the clause.

Consider the following thematic and transitivity structures of an example from Abkhaz taken from Matthiessen (1995: 370):

(10) Abkhaz "head marking": Process marked
a-xàc'a a-ph°es a-s°q°e ø-lè-y-te-yt.
'The man gave the woman the book.'

the-man	the-woman	the-book	it:to-her-he-gave
Theme	Theme	Theme	Rheme
Participant	Participant	Participant	Part.part.part.Process

And compare it with the following example from French taken from Caffarel (2006*a*: 176):

(11) Moi, ce tarif-là, je ne l'ai pas.
'**This rate, I** don't have it.'

Moi	ce tarif-là	je ne l'ai pas
Me	this rate	I don't it have
Absolute Theme 1	Absolute Theme 2	Rheme
		Medium.Range.Process

Although these two examples are very similar in terms of their realization of thematic and experiential structures, from an SFL perspective typological generalizations are not based on structure and do not attempt to categorize a whole language:

> One consequence of the multidimensional theory of language is the realization that languages are far too complex to be typologized as unified phenomena: typology has to be typology of particular systems (such as TENSE/ASPECT systems), not typology of whole languages as was done traditionally when languages were typologized as analytic versus agglutinative versus fusional versus polysynthetic (see Halliday 1966: 166–8).
>
> (Caffarel et al. 2004: 4)

When exploring a language from a typological perspective, it is also important to look at discourse, as the potential "agglutination" of participants to the process in French through cliticization is very much related to their textual statuses in discourse, whether they are introduced as new or given, and the mode of discourse, whether spoken or written.

If we look at Theme patterns in Text 2 on the beginning of the bombings in Afghanistan, we will not find any absolute Themes or participants cliticized to the verbal group for instance:

Text 2

Afghanistan: Deuxième vague de bombardements

Theme analysis: Unmarked Themes are underlined and textual Themes are in bold

(1) <u>Une deuxième vague de bombardements</u> a frappé Kaboul, tard dans la soirée de dimanche.

(2) <u>Les premières opérations anglo-américaines</u> ont commencé en Afghanistan à 18h15 (heure française).

(3a) <u>George W. Bush</u> a confirmé, dans un discours à la nation dimanche à 18h50,

(3b) **que** <u>les États-Unis et la Grande-Bretagne</u> avaient entamé des frappes militaires "ciblées".

(4) <u>Les forces armées anglo-américaines</u> ont lancé une première offensive sur Kaboul, la capitale afghane, Jalalabad, Kandahar et Hérat.

(5) <u>Le porte-parole du ministère pakistanais des affaires étrangères</u> a regretté [[que les efforts diplomatiques n'aient pu permettre de convaincre les dirigeants talibans de "répondre aux demandes internationales"]].

(6) <u>Les talibans</u> ont qualifié les bombardements "d'acte terroriste".

(7a) <u>Ben Laden</u>, dans un message préenregistré [[diffusé dimanche soir par Al-Jazira TV]], se dit prêt à "la confrontation"

(7b) **et** affirme

(7c) **qu'** <u>un "groupe de musulmans"</u> a bien commis les attentats du 11 septembre.

The first striking pattern is that we have only unmarked Themes (underlined in the text), that is, the Subject of each clause is also the Theme of that clause. This has the effect of making Actor, Sayer, and Senser involved in the Process textually prominent. The text is about participants associated with a particular event: who does what, who says what, and who feels what. However, as we saw in the transitivity analysis, the text tries to remain factual and "non-judgmental". The "neutral" stance of the text is foregrounded by having each of the participants involved in the event, the Anglo-Americans, George Bush, Pakistan's authorities, the Taliban, and Bin Laden, functioning as Themes. It is also important to note that there are no Interpersonal Themes and very few Textual Themes, the text simply listing the facts rather than trying to explain them logically. The thematic progression is mainly constant, that is, the text introduces the participants one by one and reports what they did, said, or thought. The almost total lack of linear progression where the Rheme of the preceding clause becomes the Theme of the following clause again indicates that the text is not trying to explain the situation but is simply relating the facts.

This brief overview of the metafunctional organization of the French clause has illustrated how grammar makes meaning through metafunctional diversification and grammatical metaphor. The analysis of Text 2 also illustrated how texts have meanings beyond their first-order meanings, that is, symbolic or second-order meanings, created through consistency of foregrounding of grammatical patterns.

31.5 Conclusion

This chapter focused essentially on the power of grammar within SFL as a tool for linguistic analysis and the interpretation of meanings, first- and second-order, in any text type. Discourse analysis is one of the potential applications of SFL but there are many others. Its strong orientation toward meaning has made it useful in many applied contexts, such as education (e.g., Christie and Martin 2006; Jones 2008), the study of literature (e.g., Butt 1988; caffarel 2004), translation (e.g., Steiner and Yallop 2001), computational implementations (e.g., Teich 1999), artificial intelligence (e.g., Kobayashi et al. 2002), the study of multimodal texts (e.g., O'Halloran 2006) and other semiotic systems (e.g., Kress and Leeuwen 2006), psychiatry (e.g., Fine 2006) and language disorders (e.g., Armstrong 2001). For an overview of various other applications of SFL, see Hasan et al. (2005, 2007).

This page intentionally left blank

CHAPTER 32

USAGE-BASED THEORY

JOAN L. BYBEE
CLAY BECKNER

32.1 Statement of goals of the theory

Usage-based theory takes language to be an embodied and social human behavior and seeks explanations in that context. As the name indicates, this theoretical perspective incorporates the basic insight that usage has an effect on linguistic structure. It thus contrasts with the generative paradigm's focus on competence to the exclusion of performance and rather looks to evidence from usage for the understanding of the cognitive organization of language. Thus usage patterns, frequency of occurrence, variation, and change are all taken to provide direct evidence about cognitive representation. No relevant methods for gaining evidence about language are excluded; studies of corpora, large and small, diachronic data, psycholinguistic experiments, cross-linguistic comparison and child language development all provide essential data for constructing a comprehensive theory of language.

32.2 Background

Usage-based theory has its source in a confluence of a variety of research perspectives which consider the effect that usage might have on linguistic representation.[1] One practice that unites many of these researchers is a methodological one: it is common now to address theoretical issues through the examination of bodies of naturally-occurring language use. This practice has been in place for decades in the work of those who examine the use of grammar in discourse with an eye toward determining how discourse use shapes grammar, notably Givón, Thompson, Hopper, and DuBois (e.g., DuBois 1985; Givón 1979; Hopper and Thompson 1980; Ono et al. 2000; Thompson and Hopper 2001). In addition, researchers in sociolinguistic variation, such as Labov, Sankoff, and Poplack (e.g., Labov 1972; Poplack 2001; Poplack and Tagliamonte 1999, 2001; Sankoff and Brown 1976) have always relied on natural discourse to study the inherent variation in language use.

Usage and text-based research, always central to traditional historical linguistics, is especially emphasized in functionalist work on grammaticalization, e.g., Bybee (2003*a*, 2003*b*), Hopper and Traugott (2003), and Poplack and Tagliamonte (1999). In fact, the study of grammaticalization has played a central role in emphasizing the point that both grammatical meaning and grammatical form come into being through repeated instances of language use (see section 32.7.3).

Of course, one major impetus for the shift to analysis of natural language use is the recent availability of large electronic corpora and means for accessing particular items and patterns in such corpora. Through the work of corpus linguists, such as John Sinclair (1991), computational linguists, such as Dan Jurafsky and colleagues (e.g., Jurafsky et al. 2001; Gregory et al. 1999), and those who are proposing probabilistic or stochastic grammar, such as Janet Pierrehumbert (e.g., 2001), Rens Bod (1998; this volume), access to the nature and range of experience an average speaker has with language is now within our grasp. Studies of words, phrases, and constructions in such large corpora present a varying topography of distribution and frequency that can be quite different from what our intuitions have suggested. In addition, the use of large corpora for phonetic analysis provides a better understanding of the role of token frequency as well as specific words and collocations in phonetic variation.

At the same time a compatible view of language acquisition has been developing. The uneven distribution of words and constructions in speech to children is mirrored somewhat in the course of acquisition: children often produce their first instances of grammatical constructions only in the context of specific lexical items

[1] The term "usage-based" comes from Langacker (1987*b*; 1988); see Barlow and Kemmer (2000).

and later generalize them to other lexical items, leading eventually to productive use by the child; see Tomasello, Lieven, and their colleagues (e.g., Lieven et al. 2003; Tomasello 2003; Savage et al., 2003; Dąbrowska and Lieven 2005).

32.3 Domain-general processes

Usage-based theory postulates that the units and structure of language emerge out of specific communicative events (section 32.4), and strives to avoid relying on innate knowledge specific to the domain of language. A usage-based model thus takes as its null hypothesis the view that language is an extension of other cognitive domains. Elman and Bates (1997: 1,180) write that "language evolved through quantitative changes in social, perceptual, and cognitive abilities, including statistical learning, that exist in other species. These abilities have been recruited for language, but they continue to do nonlinguistic work (that is, they have kept their 'day jobs')". Along these lines, usage-based theory seeks to derive the mechanisms of language from more general and basic capacities of the human brain, including sequential and statistical learning, chunking, and categorization.

32.3.1 Repetition, chunking, and knowledge of usage

A general characteristic of cognition is that repetition of an activity has a cumulative effect on future behavior. In the domain of motor skills—as in learning to ride a bicycle, for instance—an initially deliberate, difficult task can be automatized with practice, eventually becoming an unconscious routine (McCrone 1999). Repetition of an activity causes us to develop "procedural knowledge", that is, implicit knowledge about *how* to do something, in contrast with explicit, declarative knowledge (Anderson 1993). Across domains, learning involves a feedback loop: the human cognitive system produces actions while also monitoring and updating itself on the basis of these actions. With respect to motor activity and other cognitive processes, experiments show that repeatedly engaging in a task leads to the formation of a representation of that process in long-term memory (Shadmehr and Brashers-Krug 1997), and "chunks" the process into useful sub-routines (Simon 1974; Graybiel 1998; Sakai et al. 2004).

In general, it seems that our cognitive systems track any behaviors that keep occurring, improving performance by rendering the activity into chunks that make processing more efficient (Haiman 1994). This principle seems to hold across domains, both for events presented in isolation, and for multiple events that co-occur or occur in sequence. People are quite good at learning when two (or more) events tend to co-occur, or when one event tends to predict another; such abilities are indeed shared by animals other than humans (Kelly and Martin 1994; Bush 2001). A variety of artificial grammar studies in recent years (e.g., Saffran et al. 1996) have demonstrated that people are strikingly skilled at detecting patterns, and inferring units of co-occurrence, based on transitional probabilities, on the basis of relatively little input. Both small children and adults learn such patterns relatively automatically, whether the input consists of language-like syllables, or unlanguage-like stimuli like tones or shapes (Saffran et al 1999; Fiser and Aslin 2002). This pattern detection is a domain-general process of the human mind: we pursue it without conscious effort, and whether or not there is a communicative reason to do so.

Experience thus has an ongoing effect on mental representation. With respect to this broad principle, usage-based theory holds that there is no reason to claim that language is different from any other cognitive domain. A speaker's knowledge of language incorporates a large body of implicit, procedural knowledge, including knowledge of frequency and statistical patterns (Bybee 1998; 2002*a*; Bybee and Hopper 2001*a*; Gahl and Garnsey 2004). A usage-based view holds, further, that there is little reason to claim that knowledge gathered from ongoing experience is fundamentally separate from core knowledge of the language (e.g., "competence" or I-language; see Lightfoot 2006; Newmeyer 2003; 2006 as examples of the generative view).

It would in fact be surprising if experiential knowledge needed, for some reason, to be quarantined from the rest of linguistic knowledge. Anderson (1978: 273) observes that "well-designed systems tend to have special representations for the kinds of information they have to process frequently",[2] drawing examples from visual and auditory processing, and human-designed systems in computer science. The online demands of processing language, both in perception and production, are not trivial. Compared with a static generative model, a system in which mental representations are updated on the basis of incoming information (e.g., a usage-based system) would seem to be more likely to operate smoothly in the face of such demands, and is more in line with what we know about other areas of cognition.

As we will see below (sections 32.5, 32.6, and 32.7), there are further reasons to believe that knowledge of usage is a core part of linguistic knowledge, given that

[2] We wish to be cautious in interpreting the term "well-designed" in the present context. We certainly make no claim that language (or any other cognitive capacity) is externally designed, nor that it is maximally optimized. Our perspective instead is that language is a self-organizing system (Camazine et al. 2001) that exhibits certain apparent "design features". One of these emergent design features is that the grammar is rendered more efficient by encoding frequency information, resulting from a domain-general pattern in which mental representations are updated rather than remaining static.

procedural knowledge is implicated in lexicon and grammar change. Frequency along with other usage-based factors must be incorporated into the grammar, because repetition is necessary to the operation of the common mechanisms of language change (Haiman 1994; Bybee 2006*a*).

32.3.2 Categorization

Categorization represents another domain-general capacity which is of central importance in usage-based theory. We have noted that mental representations are continually shaped by the repetition of events, but for repetition to be recognized, people must sometimes consider two events to be "the same" despite some differences (Haiman 1997). Across domains, categorization allows us to map continuously varied input into "equivalence classes" in some context, on the basis of shared properties (Bruner et al. 1956; Pierrehumbert 2001). For instance, people can quickly learn to classify visual stimuli on the basis of examples they are exposed to in an experimental setting (e.g., Posner and Keele 1968, Medin and Schaffer 1978, Notman et al. 2005), and rapid visual classification of certain complex scenes can occur seemingly effortlessly even without focused attention (Li et al. 2002).

Although we are indeed able to group together input having varied properties, category membership is a gradient, rather than an absolute, phenomenon. There is little evidence for the classical model of categorization, in which categories are defined by necessary and sufficient conditions (Rosch 1978). One category is not sharply defined from the next, but rather the boundaries are gradient (Labov 1973). Moreover, in contrast with the classical theory, categories have an internal structure, and some members are "better members" than others. For instance, with respect to identifying members of the conceptual category "fruit", American participants find that especially good examples are *apple, orange*, and *banana*, but less central examples might be *watermelon, raspberry*, and *mango*. Such internal category structures become evident via a variety of converging methodologies, including typicality ratings, response times for classification tasks, and the order in which items are listed in a production study (e.g., Battig and Montague 1969; Rosch 1975; Van Overschelde et al. 2004).

In one framework, these category-internal structures are said to derive from relations to a category prototype that encapsulates a central tendency (Rosch 1978). Degrees of category membership then extend outward from the prototype in a network of partially-shared features, resulting in a category with a "family resemblance" structure (Rosch and Mervis 1975). Yet further study has led to the finding that our knowledge of categories cannot just be based on an abstract summary but must include representations for individually experienced tokens.[3] For example,

[3] A full discussion of the merits of prototype vs. exemplar models is not possible in the present chapter; see Medin and Schaffer (1978) and Ross and Makin (1999). The evidence indicates that

Posner and Keele (1968) studied subjects' classification of visual dot patterns under different training conditions in which the central tendency for the category was held constant. Subjects who learned the category on the basis of a more variable training sample were better at classifying noisier variations on the category, compared with subjects who learned under a less variable condition. Since both groups were presumed to have the same category prototype (and the low-variability group should have learned that prototype better), this result is not expected unless learners maintain knowledge about individual exemplars. Another bit of evidence for extensive exemplar storage comes from the finding that people are aware when certain features tend to co-occur *within* a particular category. For instance, people implicitly know that if a bird sings, it is much more likely to be a small bird than a large bird (Malt and Smith 1984). This detailed, intra-category knowledge is not explainable if people only represent the category using an abstract "bird" prototype, while discarding knowledge of individual exemplars.

Evidence such as the foregoing implies that we do not reduce categories to minimal abstractions but rather maintain representations for both coincidental features and highly predictable traits for the category. Moreover, it seems that people retain memories of individual members of a category, since the structure of categories is known to be influenced by the frequency with which particular items are experienced (Nosofsky 1988). All these findings are crucial in usage-based theory, which holds that in language and other domains, specific instances of learning are retained in memory alongside the generalizations that gradually emerge from them (see section 32.4). Moreover, usage-based theory maintains that linguistic categories are just like categories from any other cognitive domain; there is a rich, item-specific internal structure to categories in phonology (Miller 1994), morphology (Bybee and Moder 1983), and grammatical constructions (Goldberg and Giudice 2005; Bybee and Eddington 2006). As we argue below (section 32.4.2), linguistic units are gradient categories that have no fixed properties but rather are formed on the basis of experienced tokens.

32.4 The formal apparatus: Exemplars, networks, and constructions

In contrast to earlier theories that assume limited memory capacity and thus attempt to separate the predictable from the idiosyncratic by representing the latter

exemplar models can fully account for prototype effects, and can explain some findings that are not predicted by prototype models.

in the lexicon and the former in rules, usage-based theory takes a nonreductive, non-minimalist approach to linguistic representation (Bolinger 1976; Langacker 1987*b*, 1988; Booij, this volume). We take into account the extensive evidence that speakers maintain "rich memory representations" in which experiences with language in all their glorious detail are stored in exemplars (Tenpenny 1995; Goldinger 1996; 2000; K. Johnson 1997). In addition to specific exemplars of experienced language, categorization of these exemplars provides more abstract generalizations or schemas. While generative theories emphasize the abstractions, in the current framework we are interested in how the specific experiences speakers have with language combine to yield more general patterns, and how the specific and general interact in acquisition, processing, and language change.

In an exemplar model every token of use impacts cognitive representation. In phonetic perception and decoding, if an input token is the same as an existing exemplar, it is mapped onto that exemplar, strengthening it. If it is not similar enough for a mapping to an existing exemplar, a new exemplar is established, positioned in a metaphorical space close to similar exemplars (Bybee 2001*a*; Pierrehumbert 2001). Thus for every word in a speaker's lexicon, there is a cloud or cluster of phonetic exemplars representing all the phonetic variants of word with information about their linguistic context and further indexes to the social context (Foulkes and Docherty 2006). In speaking, one of these exemplars is chosen for production (Pierrehumbert 2001; 2002). The meaning of the word is also represented by a cluster of exemplars which represent the context and meaning for each token of a word. It is proposed that memory for linguistic objects is the same as for non-linguistic objects, which means that memories can also decay. Particular exemplars that are marginal and not reinforced may be lost, keeping word (and other) categories centered in both their form and meaning (Pierrehumbert 2002; Wedel 2006).

Although every token of experience does affect the system in an exemplar model, we should also take note that not every token produces sweeping change! In fact, as a general rule, in an exemplar model new input either further reinforces an existing pattern, or produces a relatively small change in the system's probabilities. Exemplar models thus provide a framework in which usage-based theory can explain both diachronic and synchronic regularities in language—a necessity for an adequate linguistic theory since we must account for the fact that communities exhibit quasi-stability in speech conventions over time, in addition to the fact that languages *do* change in certain ways. On the one hand, with experience speakers accrue a store of exemplars which may lead to progressively advanced entrenchment via an ongoing production-perception feedback loop. In a population of speakers, stability may be further encouraged by the collective weight of accrued conventions multiplied out over an entire speech community. On the other hand, system equilibrium is anything but inevitable in an exemplar model. New exemplars (involving any combination of phonological, morphosyntactic, or semantic-pragmatic traits) may

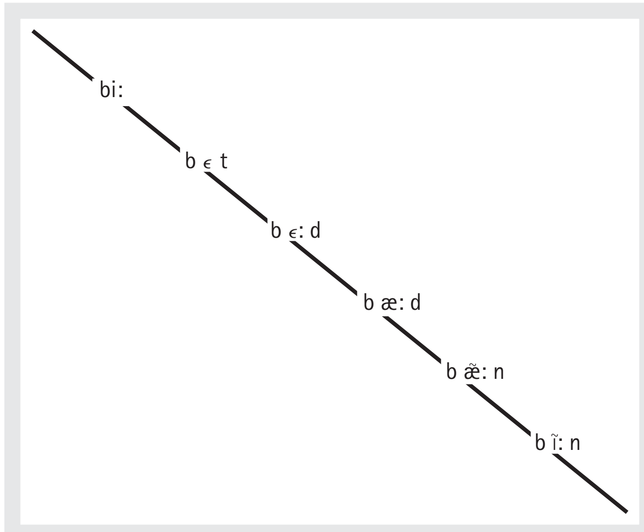


Figure 32.1. Lexical connections for the [b] in *bee, bet, bed, bad, ban, bin*

filter into the system via dialect contact or via the internal mechanisms of reduction, chunking, categorization, analogy and inference discussed below.

32.4.1 Networks

Similarities among words and even longer strings are represented in networks. Through these networks, units of language on various levels emerge. Networks arise through categorization; when tokens of linguistic experience share properties with established exemplars, but also differ in some way, then their shared properties are linked or located close by in mental "space". In the diagrams we will use for illustration—following Bybee (1985*a*) such links are shown as lines—solid lines for identity and broken lines for similarity.[4] These links establish units smaller than the word. Figure 32.1 shows phonological connections; Figure 32.2 shows parallel phonological connections and semantic connections that occur across a number of items; in this case we can speak of affixes emerging from the categorization. Figure 32.3 shows how the internal structure of a complex word emerges through the comparison with related words.

Considerable evidence has been presented in recent work to show that multiword phrases can also be stored in memory. In the case of idioms, which have meanings

[4] The schemas of Booij (this volume) may be interpreted as generalizations that capture the types of relationships expressed by network diagrams. In the network diagram convention, generalizations are understood to be implicit and emergent from the network, rather than being overtly notated in the diagram. (See also Bybee 2001*a*: 22.)

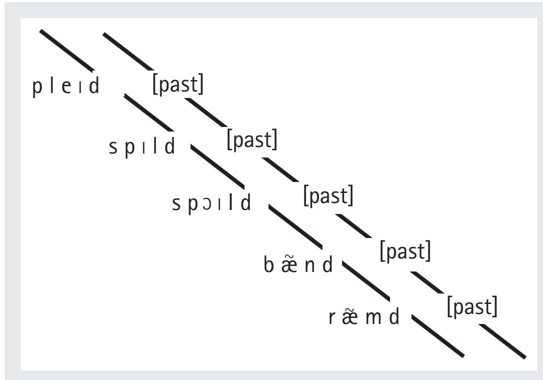


Figure 32.2. Phonological and semantic connections yield Past in *played, spilled, spoiled, banned, rammed*

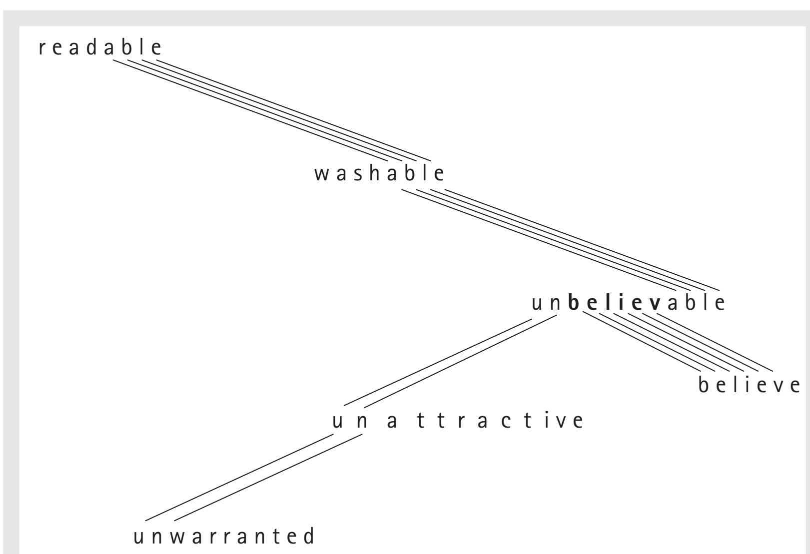


Figure 32.3. The internal structure of *unbelievable* emerges from connections to related words

that are not transparently compositional, such storage is necessary by traditional standards. However, the existence of other collocations, commonly referred to as "prefabs", which do not necessarily have any idiosyncrasies of meaning or form but are conventionalized expressions and known to speakers as such, argue for more extensive storage of multiword sequences (Pawley and Syder 1983; Erman and Warren 2000). Thus for instance, while *pull strings* as in *he pulled strings*

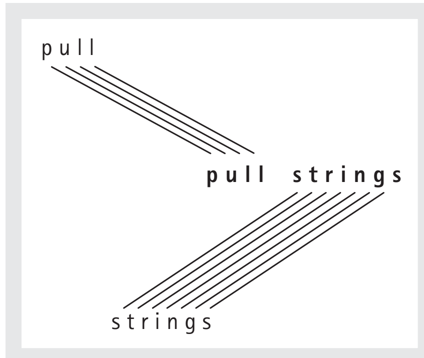


Figure 32.4. The connections between an idiom and its component words

to get that job has a metaphorical meaning, the phrases *for some reason* or *dark night* are transparently compositional in form and meaning and yet represent the conventional way of expressing certain notions. Knowledge about the conventionality of all these sequences must be represented somehow in the grammar, since fluent speakers do not produce (or accept) the full range of utterances permitted by combinatoric syntactic rules. (Compare the non-conventionalized and rather awkward *by some reason, for some cause*, and *black night*.) In the case of idioms and prefabs, their representation in memory does not preclude the speaker knowing what the constituent words are, nor does it preclude access to their meanings and other uses (Nunberg et al. 1994). From a usage-based perspective, there is no need to choose between storage of an unanalyzable unit and compositional assembly, since speakers may in fact have a rich and multifaceted representation for a sequence. A network representation is quite appropriate as it allows access to the sequence as a whole, while maintaining the links that identify the component parts, as illustrated in Figure 32.4.

32.4.2 Units and levels as emergent

All of the units of language—segments, phonemes, morphemes, words, phrases, constituents—can be arrived at by the simple categorization process described above. They do not have to be postulated as a part of the innate universal grammar because they can be arrived at by speakers based on the input and the domain-general process of categorization. The strings of linguistic material that are experienced by the learner are stored in memory (perhaps imperfectly at first) and the brain automatically searches for similarity among such stored experiences, placing them in networks based on these similarities. Whatever repeated units occur in the experience of the learner will emerge in the networks. Thus if the child hears *ice cream* in different linguistic contexts, such as *I like ice cream* and *do you want some ice cream*, the string *ice cream* will emerge from comparisons of similarity. Note that the speaker/learner is also registering in memory the extra-linguistic contexts in which the linguistic material occurs; in this way semantic and pragmatic representations are also set up. Thus, given certain constraints (such as token and type frequency, see section 32.5), the learner will find the regularities that occur in the input.

This theory raises the question of why languages have units such as segments, affixes, stems, words, and constructions. Rather than postulating such units as givens (innate in the language learner), usage-based theory leaves open the possibility of actually explaining why languages have such units and how they differ across languages (see Lindblom et al. 1984 for an early expression of this view). This explanation will look to diachronic processes to explain current language states. It is worth noting that the postulation of linguistic units as innate universals does not stand up well given the real facts of language, which show, as we will see in the next section, that distinctions between unit types are blurred by both gradience and variability. These facts indicate that dynamic processes rather than static universals are creating regularities.

32.4.3 Gradience

Gradience refers to the fact that the boundaries of many categories of grammar are difficult to distinguish, usually because change occurs over time in a gradual way, moving an element along a continuum from one category to another. Continua such as those between function words and affixes, between regular and irregular patterns, and between productive and unproductive constructions illustrate this gradience.

To demonstrate how the exemplar cum network representation allows for gradience, let us consider some examples, starting with morphemes. Morphemes are traditionally considered to be form–meaning pairings, but problems with the premise that all strings are exhaustively dividable into morphemes have been long noted in the literature. Two types of problems occur. First, dividing words into morphemes sometimes yield leftover bits that are not themselves morphemes. Dubbed "cranberry" morphs by structural linguists, *cran* is one of them because *berry* is obviously a morpheme, but what is *cran*? Other examples are the *Tues*- and *Wednes*- of the days of the week, where the morpheme -*day* is recognizable but the front part of the word is not. These are not problematic for the network model because whole words are stored in the lexicon and there is no requirement that all parts of a word be connected to some other word. Thus *berry* and *day* can have their links while the other parts of the word can be unconnected (Bybee 1988*a*).

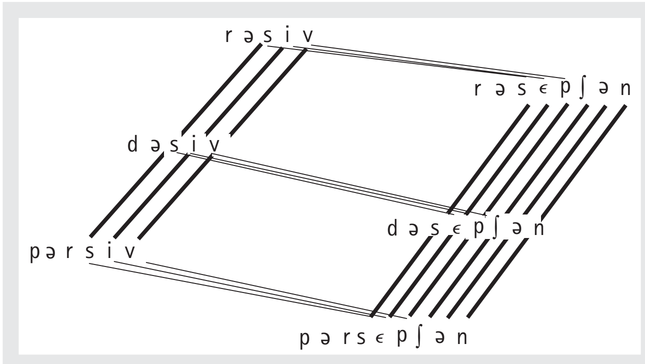


Figure 32.5. How recurring word parts are recognized

Second, some parts of words appear to be repeated across words, but they lack a discernible meaning. For instance, *-ceive* (with its alternate, *-cep-*) occurs in words such as *receive, reception; deceive, deception; conceive, conception; perceive, perception*, revealing a minor pattern of which most speakers probably are aware, yet this Latin stem has no meaning in English. Thus the category of "morpheme" shows gradience in that formal bits of language exist that are recognizable but not fully meaningful. In the network, the phonological associations are made, as in Figure 32.5, but no semantic connections are made, except those between the base and the derived form, and those relating to the categories of verb and noun.

Another gradient dimension in morphology ranges between regular vs. irregular morphological formations. In some theories (generative theories and Pinker's (1991) dual processing model) regulars and irregulars are thought to constitute discrete types and to be processed in totally different ways. In the usage-based model, differences arise in the way complex words are processed due to differences in token and type frequency (see section 32.5). One argument for not drawing a strict line between regulars and irregulars is that there can be substantial overlap between the two types. Thus McClelland and Patterson (2002) point out that many so-called irregular verbs in English have the *t/d* suffix of the regulars, e.g., *slept, brought, went*, etc.[5]

Other instances of gradience in grammar concern the degree of grammaticalization, which of course changes over time (see Heine and Narrog, this volume) and which gradually moves units from independent words, to function words, to affixes. The gradualness of linguistic change means that at any given moment in a synchronic grammar, there will not only be variation, but also gradience in the sense that some units will not fall squarely into the linguist's categories of word, clitic, or affix.

[5] For a presentation of the usage-based approach as applied to phonological phenomena, see Bybee (2001*a*).

32.4.4 Larger units

Much of our subsequent discussion will focus on the gradient properties of larger syntactic units—in this framework, constructions, which we will discuss more thoroughly in sections 32.6 and 32.7. Constructions are conventionalized sequences of morphemes or words that contain a position that can be filled by more than one item. Consider, for example, the expression *drive someone crazy, mad, insane, nuts, up the wall*, etc. This is a construction that contains the specific verb *drive* (in any of its inflected forms), an object pronoun, and an adjective denoting a state ranging from true insanity to extreme irritation. Such a construction can emerge from a network via exposure to specific tokens. *Drive* + object pronoun is the anchor, i.e., the most stable part of the expression, and the adjective slot is more open, though it is semantically constrained. This analysis is arrived at again by categorization based on similarity of form and meaning for *drive* and meaning only for the adjectives.

Like the other units we have discussed, constructions exhibit both gradience and variation, since they can vary in their degree of grammaticalization, productivity, schematicity and their appropriate contexts of use, as we will see below.

32.5 The role of repetition: Effects and mechanisms

We have already seen that exemplar models register variation and change while it is ongoing. In addition, exemplars are strengthened by repetition, so that frequency is naturally represented in cognition. The network of connections is also sensitive to frequency of use. In this section we consider both TOKEN FREQUENCY—the number of times an item or string occurs in running text, and TYPE FREQUENCY—the number of distinct items that can be used in a pattern.

32.5.1 The reducing effect of high-token frequency

An extensive body of literature has shown that high-frequency words and phrases undergo phonetic reduction at a faster rate than low- and mid-frequency sequences (Schuchardt 1885; Fidelholtz 1975; Hooper 1976; Bybee and Scheibman 1999; Bybee 2000*b*; 2001*a*). This REDUCING EFFECT applies to phrases of extreme high frequency such as *I don't know*, which shows the highest rate of *don't* reduction (Bybee and Scheibman 1999), and also to words of all frequency levels undergoing gradual sound change, such as English final t/d deletion or Spanish [ð] deletion, both of which affect high-frequency words earlier than low-frequency words (Bybee 2001*a*; 2002*b*; Gregory et al. 1999). This effect of repetition is the result of the domain-general processes discussed in section 32.3.1. Words and phrases represent "chunks" of neuromotor behavior. With repetition, their execution becomes more fluent as the articulatory gestures involved reduce in magnitude and overlap adjacent gestures (Pagliuca and Mowrey 1987; Browman and Goldstein 1992).

32.5.2 Entrenchment and autonomy: The conserving effect of high token frequency

Alongside the Reducing Effect, words and phrases with high-token frequency are also subject to the Conserving Effect, meaning that high-frequency items are more resistant to reformations based on productive patterns in the language. These two effects may seem paradoxical at first glance, but they are caused by two different cognitive mechanisms which respond to token frequency: in addition to increasing fluency, high token frequency has the effect of strengthening memory representations. This strength is reflected in easier lexical access and, in complex words and strings, resistance to reformation. For any given string that consists of more than one meaningful element, there can be at least two ways of accessing it: either as a single unit or as a set of units that are then combined into a whole. For instance, the word *insane* can either be accessed as a unit, or it can be built up by combining the prefix *in-* with the stem *sane* (Hay 2001).[6] The higher the token frequency of the sequence, the more likely it will be to be stored and accessed whole (Bybee 1985*a*; but see also Hay 2001). Thus high-frequency sequences are more entrenched in their morpho-syntactic structure and therefore resist change on the basis of more productive patterns. Among English irregular verbs the low-frequency verbs are more likely to regularize (*weep, weeped*) while the high-frequency verbs maintain their irregularity (*keep, kept*). The reason is that frequency strengthens the memory representations of words or phrases, making them easier to access whole and thus less likely to be subject to reformation on the basis of more productive patterns (Hooper 1976; Bybee 1985*a*). This effect applies to syntactic sequences as well, allowing higher-frequency exemplars to maintain a more conservative structure (Bybee and Thompson 1997). In section 32.7.2 we discuss several examples that show the maintenance of the older constructions in high-frequency contexts.

As we said in our discussion of networks, chunks that are stored whole can also maintain their associations with other instances of their component parts. In cases of extreme high frequency, however, a morphologically complex form (or string of

[6] Building up strings of morphemes of words vs. accessing them already assembled do not actually constitute two mutually exclusive means of access; rather they represent two poles of a gradient, which is the extent to which the separate components of a string are activated when the whole string is activated.

words) can lose its internal structure and become autonomous from etymologically related forms (Bybee 1985*a*). This can be seen, for example, in the way that words with derivational affixes become less transparently related to their base forms as they become more frequent (Bybee 1985*a*; Hay 2001). Hay (2001) argues that the semantic opacity of words such as *dislocate*, etc. is due to the fact that their complex forms are more frequent than the bases from which they were originally derived. The effect applies to inflection only in cases of extreme high frequency where it leads to suppletion. Thus *went* was formerly the past tense of *wend* but (for unknown reasons) it increased in frequency and moved away from *wend*, joining *go* to become the past tense of that verb. This effect also applies in grammaticalization when sequences that are originally complex (such as *be going to*) lose their semantic and syntactic transparency and move away from other instances of the Progressive, *go*, and *to*.

32.5.3 Type frequency, schematicity, and productivity

Type frequency is a property of patterns or constructions and refers to the number of distinct items that can occur in the open slot of a construction or the number of items that exemplify a pattern, such as a phonotactic sequence. For instance, the regular English Past Tense inflection with *-ed* applies to thousands of verbs and thus has a very high type frequency. In contrast, the vowel-change pattern exemplified by *string, strung; fling, flung; stink, stunk* applies to some eighteen English verbs and thus has a lower type frequency. Taking a phonotactic example, the word-initial sequence *sp-*, as in *spark, spot, spin*, etc., has a much higher type frequency that the cluster *sf-*, as in *sphinx* and *sphere*.

Type frequency is the main factor that determines the degree of productivity of a construction (Guillaume 1973 [1927]; MacWhinney 1978; Bybee 1985*a*). That is, patterns or constructions that apply to a high number of distinct items also tend to be highly applicable to new items. In determining productivity, however, factors other than type frequency must also be taken into account: often the member items that occur with a construction must also belong to certain phonological or semantic categories. For instance, the verbs of the *string, strung* class must end in a nasal or a velar (Bybee and Moder 1983). The open slots in constructions are often semantically restricted, as the adjectives that can be used in the construction [X drives me (or someone) ADJ] (as in *it drives me mad, it drives me crazy*) must suggest some degree of insanity, either literally or figuratively (Boas 2003). Thus productivity is a matter of degree, determined by an interaction of type frequency with schematicity—the degree to which the category is open or restricted.

The contribution of type frequency to productivity comes about when a construction is experienced with different items occupying a position, which enables the parsing of the construction (Hay and Baayen 2002). If *happiness* is learned by someone who knows no related words, there is no way to infer that it has two morphemes. If *happy* is also learned, then the learner could hypothesize that -*ness* is a suffix, but only if it occurs on other adjectives would its status as a suffix become established. Thus a certain degree of type frequency is needed to uncover the structure of words and phrases. In addition, a higher type frequency also gives a construction a stronger representation, making it more available or accessible for novel uses. Schematicity contributes to productivity in that highly schematic categories are more easily extended to new items. Since there are no phonological or semantic restrictions on the regular English Past Tense suffix -*ed*, it is free to apply to any verb.

Thus productivity and schematicity are highly related to categorization since the application of a construction depends upon the properties of the category formed for the open position. Both types and tokens contribute to categorization. The *properties* of the types included in a category establishes its range or schematicity while the *number* of types relates to the degree of productivity of the construction referring to the category. In research into exemplar models (in which the category consists of the experienced exemplars), token frequency has been shown to influence the perception of the center of the category, as well as its boundaries (Nosofsky 1988). In phonetic categorization, high-frequency exemplars tend to be maintained while low-frequency ones are marginalized and lost (Bybee 2001*a*; Pierrehumbert 2001). In semantic categorization a similar phenomenon occurs; in a corpus and experimental study of the pairing of verbs meaning "become" with adjectives in Spanish, it was found that the high-frequency pairs served as the center of some of the most productive categories (Bybee and Eddington 2006). Similarly, Casenhiser and Goldberg (2005) show that children and adults learn a new construction faster if they are exposed to one higher-frequency token as well as several types exemplifying the construction.

32.6 Constructions: Form–meaning pairings

For the purpose of syntactic description, the usage-based model adopts constructions as the basic unit of form–meaning correspondence (Fillmore et al. 1988; Goldberg 1995; 2006; Croft 2001). We regard any conventionalized string of words or morphemes as a construction, but our focus for an understanding of syntactic productivity is on strings that include at least one schematic position—a position in which more than one word or morpheme may appear. What we regard as the grammar of a language is a collection of constructions, organized into networks by the same criteria that words are—by their formal and semantic similarity.

An important property of a grammar based on constructions is that it reflects the deep intertwining of lexical items with grammatical structure. Most constructions contain very specific lexical material, such as the verb *drive* in the *drive someone crazy* construction mentioned above, or *-ed* (and its allomorphs) in the regular Past Tense construction. In addition to having fixed linguistic material, most constructions restrict the set of lexical items that can fill the open position, as when *drive someone* ____ must contain an adjective or prepositional phrase meaning "crazy". The fact that a certain lexical item (in this case *crazy* for American English) occurs more often in this slot than any other lexical item is recorded in the exemplar representation as important information for the category of items occurring there. In other words, in an exemplar model constructions are not abstract grammatical patterns but rather they are sets of experienced exemplars arranged in cognitive space to reflect their similarity in form and meaning.

Consider in more detail the *drive someone crazy* construction, as studied by Boas (2003) (cf. a set of "become" constructions in Spanish as analyzed by Bybee and Eddington 2006). This construction uses the verb *drive* with an adjective or prepositional phrase expressing a meaning such as "drive crazy". Particular tokens found in the British National Corpus (BNC) include:

(1) It drives me crazy.

(2) He was going to drive her crazy if she wasn't careful.

(3) That old thing, it's just driving us crazy.

(4) They drive you mad.

(5) The death of his wife the following year drove him mad.

(6) It drove the producer mad.

(7) A couple of channels that used to drive her up the wall. . .

(8) This room drives me up the wall.

For illustration, the eight tokens represented above could each be considered exemplars which are grouped together with their identical parts mapped onto one another and their schematic parts forming categories as in (9). The adjectives illustrated here are *crazy, mad*, and *up the wall*; the others that occur in the BNC are semantically related to these (see Boas 2003).

(9)

$$\textsc{subject}\ \ [\text{DRIVE}] \left\{ \begin{array}{l} \textit{me} \\ \textit{us} \\ \textit{you} \\ \textit{him} \\ \textit{her} \\ \textit{the producer} \end{array} \right\} \left\{ \begin{array}{l} \textit{mad} \\ \textit{crazy} \\ \textit{up the wall} \\ \quad \ldots \end{array} \right\}$$

The category of SUBJECT has not been represented with actual exemplars because it appears to take any NP. Presumably NP is a category that can be developed on the basis of the exemplars that occur in other constructions (Croft 2001). [DRIVE] is a notation intended to show that any inflected form of the verb *drive* may appear, in addition to any of the other auxiliary or emerging auxiliary constructions (e.g., *used to, gonna*...). The enlarged font of [DRIVE] represents the strength it acquires by occurring in all instances of the construction. *Mad* and *crazy* are similarly shown enlarged because of their high frequency in the construction. The experiencer slot is usually a pronoun, but is always animate and usually human. The final position, which can be an adjective or prepositional phrase, has a strong semantic character. Most of the fillers for this slot found in Boas' study of the BNC were synonyms with "crazy", though there were also slightly more distantly related senses such as *to desperation*, or *to suicide*. Note that the category of adjectives and prepositional phrases is not represented as an abstraction but rather by specific items, since these exemplars are retained in memory along with knowledge of their respective frequencies. Novel additions to this category are made on the basis of analogy with existing exemplars. We propose, following the evidence in Bybee and Eddington (2006), that the most frequent members of this category serve as the center of the category; not only are they more likely to be chosen for subsequent productions but they also serve more often than any others as the basis for analogy.

Most of the constructions discussed in the literature are somewhat specific, as is the one discussed here. For this reason, some researchers doubt that a construction-based account can ratchet up to a full account of syntactic phenomena. For instance, Jackendoff (2002) accepts constructions as necessary in a grammar, but in addition maintains phrase structure rules. In contrast, we are confident that there are no empirical data of morphosyntax that cannot be adequately described via constructions and networks of constructions. This is a pressing issue for further research.

As further evidence for exemplar representation of constructions, consider the fact that such representations allow the association not just of form and meaning but also of pragmatic implications and social contexts of use, which we know from studies of change as well as variation are important parts of the knowledge that speakers have about their language (Traugott and Dasher 2002; Torres-Cacoullos 2001). This topic is treated in more detail in section 32.7.

Finally, we note briefly that the construction-and-exemplar framework we have sketched out in this section further fits into a unified usage-based model that incorporates language acquisition. Recent usage-based accounts of acquisition (for instance, Tomasello 2003; Goldberg 2006) view constructions as a basic building block in learning a language, as children learn verbs in the context of particular sequences that pair form and function. Children first comprehend (Roberts 1983; Akhtar and Tomasello 1997) and produce (Tomasello 2000) particular verbs only in highly specific contexts, gradually expanding on these to arrive at more abstract syntactic representations.[7] Dense corpus studies of child–parent interactions also find that children are very conservative learners who are guided by particular exemplars they have learned. Lieven et al. (2003) found that a majority of the utterances (63%) by a two-year-old child consisted of exact repetitions of utterances that occurred earlier in the corpus. Moreover, among the utterances that *were* novel, 74% needed only a single operation (such as adding or removing a word) to match a particular previous utterance, or even a whole class of related utterances that permit a variable slot. The overall picture that emerges is that language learners slowly generalize item-specific sequences to permit open slots, progressively linking these constructions in a network and allowing different constructions to be combined systematically (Tomasello 2000; Dąbrowska and Lieven 2005). Within usage-based theory, there is no need to assume that knowledge about particular items is purged from memory as soon as the language learner forms generalizations (Langacker 1987*b*). Indeed, we have evidence that adult speakers maintain detailed knowledge of the internal structure of constructional categories, including a sensitivity to frequency (Bybee and Eddington 2006).

32.7 Variation and change

As we have seen, gradience and variability are built into an exemplar model: cognitive representations will reflect any new variants or ongoing changes in the distribution and frequency of variants. In this section we show that exemplar representation of constructions also provides a means to understand the creation of new constructions, the competition between constructions, and the grammaticalization of constructions.

[7] For additional evidence regarding the item-based nature of syntactic acquisition, see Lieven et al. (1997), Wilson (2003), Savage et al. (2003), and Dąbrowska and Lieven (2005), among others.

32.7.1 New constructions arise from specific exemplars of established constructions

The search for explanations for grammar in general and specific constructions in particular takes a diachronic perspective in this framework. If we want to know why a language has a particular feature, it is instructive to examine how it acquired that feature (Bybee 1988*a*). Thus we can take specific constructions and ask how they achieved that status in a particular language.

Consider a construction studied by Kay and Fillmore (1999) and C. Johnson (1997); they call it the WXDY? construction. It is exemplified in the famous joke, shown in (10) (also discussed in Bybee 2006*a*):

(10) Diner: Waiter, *what's this fly doing in my soup?*
Waiter: Why, madam, I believe that's the backstroke.
(From Kay and Fillmore 1999)

The joke shows the ambiguity of the sequence in italics. The usual interpretation of "what is X doing Y?" is one of surprise at incongruity accompanied by more than a hint of disapproval. Because it is syntactically indistinct from the construction from which it arose—a *what* question with *do* in the progressive—it gives the clever waiter license to interpret it as a literal question about what the fly is doing.

Interestingly, there is nothing in the form which explicitly suggests a meaning of incongruity, but the strong implication is nonetheless there. We can ask, then, how did an ordinary Wh-question with *doing* and a locative phrase acquire these implications? The answer must be that these implications arise from language use in context. The question of *what are you doing?* itself often has negative connotations. In a phone conversation, one may legitimately ask an addressee *what are you doing?*, but in a face-to-face situation the answer to the literal question should be available via visual inspection. Thus the question implies that the speaker wants some explanation not just of what the addressee is doing but *why* she or he is doing it. Similarly when this construction has a locative element, as in (11), there is the possibility of ambiguity, but the first reading is probably more common.

(11) What are you doing with that knife = 'why do you have that knife?'
or the literal meaning = 'what are you doing with it?'

The implication of disapproval, which is a subjective interpretation made in context, must have come from multiple instances of use with this negative nuance. As we have pointed out earlier, each exemplar of a morphosyntactic construction includes information about the contexts of use and this would include the inferences made in this context. We know from studies of grammaticalization that inferences can become part of the meaning of a construction (Traugott 1989; see section 32.7.3). The only way this could happen would be if language users were recording in memory the inferences in each situation and, at a point at which certain inferences become strong in certain contexts, they become part of the meaning of a construction.

The important point to note from this discussion is that new constructions arise out of specific exemplars of old constructions (Bybee 2003*b*; 2006*a*). This fact tells us much about how new constructions arise and it also provides evidence that cognitive representations of grammar include specific information about contexts of use of exemplars and their meaning and implications in these contexts.

32.7.2 Old and new constructions compete

Languages quite often have two or more ways of expressing the same or very similar meaning. Consider these examples: some English verbs express Past Tense by vowel changes (*blow, blew; write, wrote*, etc.) while other express the same meaning with a suffix (*chugged, hissed*); sentence negation has two alternate forms in cases where indefinites occur in the clause, for instance, *there was nothing to drink* and *there wasn't anything to drink*; English also has infinitives marked with *to* and unmarked infinitives that occur after modal auxiliaries.

Consider first the English Past Tense. We know that the ablauting process for forming the Past goes back thousands of years in Germanic, while the suffixation process is more recent. Also, it is well known that suffixation, with its high type frequency and productivity, has been gradually supplanting the ablauting process for more than a thousand years. The ablauting verbs that remain in the language are all of fairly high frequency, which is the main factor in their preservation (see section 32.5.2 above). Thus we can conclude that when older and newer constructions exist side by side in a language, it will commonly be the case that the older construction is preserved primarily in high-frequency contexts.

This principle can be applied to syntactic constructions as well. For instance the two ways that negation affects indefinite items within its scope consists of an older and a newer construction. The newer construction is the one with *not* and its contraction, as in *there wasn't anything to drink*. The older construction negates just the indefinites, as in *there was nothing to drink*. In a corpus-based study of cases where these two constructions have the same meaning and implications, Tottie (1991*b*) shows that the older (NEG-incorporation) construction is mostly used with high-frequency constructions such as existential *be* as in (12), stative *have* as in (13), and copular *be* as in (14):

(12) By the time they got to summer there was no more work to do.

(13) The Fellowship had no funds.

(14) As a nation we are not doing well enough. This is no new discovery.

The use of this type of negation with lexical verbs is much less common and tends to center around high-frequency verbs such as *know, do, give*, and *make*. The construction with *not* is much less restricted.

A third example concerns the marking on infinitives in English. Most infinitives use *to* as a marker, but after modal auxiliaries, the infinitive has no marker. Thus we contrast *I want to go* with *I can go*. The unmarked infinitive derives historically from a form with a suffix: Old English marked infinitives with the suffix *-an* and its variants. This suffix was eroded to -ən and later to -ə and then it was completely lost. At the same time, the *to* as infinitive marker had started out in purpose clauses, and was appearing in more and more constructions. However, the construction of modal auxiliary plus infinitive verb was already established in late Old English and had become quite frequent by the time the *to*-infinitive was spreading to more constructions. Because of the conserving effect of token frequency, the *to* has never been able to make its way into the modal auxiliary construction.

32.7.3 Grammaticalization of constructions requires frequency of use

Grammaticalization (see Heine and Narrog, this volume) is a central phenomenon of usage-based linguistics because it is the principal mechanism (or set of mechanisms) by which grammar is created, and it requires language use to take place. As we saw in section 32.7.1, new constructions arise out of exemplars of existing constructions. In grammaticalization, a further step is taken in that a lexical item within this construction takes on grammatical status. A recent example in the history of English is the development of the future marking periphrasis *be going to*. This developed out of a purposive construction meaning "to go somewhere to do something". It is important to note that uses of *go* in other constructions do not grammaticalize into futures. As recently as Shakespeare's time such a construction had its literal meaning. It was just one exemplar—but the most frequent exemplar—of the more general purpose construction exemplified by these sentences from Shakespeare:

(15) Don Alphonso,
With other gentlemen of good esteem,
<u>Are journeying to</u> salute the emperor
And to commend their service to his will. (*Two Gentlemen of Verona* I.3)

(16) . . .the kings
and the princes, our kindred, <u>are going to</u> see the queen's picture.
(*Winter's Tale* V.2)

Note that in both (15) and (16) the subjects are actually moving in space. In contemporary English *we're gonna see the queen's picture* can be interpreted simply as expression of future time.

Grammaticalization takes place as language is used. Grammaticalizing constructions make huge gains in token frequency and thus undergo the effects of high token frequency. As argued in Bybee (2003*b*), the changes that take place in grammaticalization are conditioned at least in part by high frequency of use. The following is a brief explanation of how frequency of use helps to condition the changes that took place in this construction. Note that all of these changes are intricately interrelated.

First, as we saw above, phonological reduction takes place when words and phrases are often repeated. Thus the increasing token frequency of *be going to* leads to the creation of a neuromotor routine that is processed as a single unit and can undergo phonological reduction to the form spelled *gonna*. Indeed, the highest frequency expression involving *be going to* is *I'm going to*, which is often produced as [aɪmənə].

Second, the autonomy of a new construction is conditioned by frequency as explained in section 32.5.2. That is, as a particular string grows more frequent, it comes to be processed as a unit rather than by its individual parts. As it is accessed more and more as a unit, it grows autonomous from the construction that originally gave rise to it. It loses its association with the purpose construction and also with the other instances of the verb *go*.

Third, the loss of the specific meaning of movement in space and addition of inferential meaning from the context also relies on frequency of use. The *be going to* construction in many contexts carried the pragmatic inference of intention, as shown in the following exchange from *Two Gentlemen of Verona* as cited in Hopper and Traugott (2003).

(17) Duke Sir Valentine, whither away so fast?
Val. Please it your grace, there is a messenger
That stays in to bear my letters to my friends,
And I am going to deliver them.
(1595, Shakespeare, *Two Gentlemen of Verona* III.i.51)

In this example, the Duke's literal question is "where are you going?" Valentine's answer does not specify location but rather intention. Interestingly, that is actually what the Duke wanted to know. The inference of intention often accompanies the use of this construction. Repeated instances such as this one make "intention" part of the meaning of the construction. The meaning and contextual implications of a construction form an exemplar cluster much as the phonetic variants do. These clusters are susceptible to the same sort of reorganization we have discussed with respect to phonetics: high-frequency semantic/pragmatic exemplars come to dominate the cluster and lower frequency exemplars may be lost, bringing about gradual semantic change.

Example (17) shows how the meaning of "intention" becomes associated with *be going to*; this interpretation is still available today. However, a further inferential change has also taken place: the expression of intention can give rise to the inference of prediction about a future event (see Bybee et al. 1994).

Finally, because items that are used together frequently come to be processed together as a unit, changes in constituency and category can take place. Thus *going to* as the constant part of this construction becomes a single unit not just phonologically but also syntactically. As the construction acquires new nuances of meaning and loses its motion sense, the following verb is taken to be the main verb. This process, known as "reanalysis", is viewed in a usage-based perspective as being gradual, that is, as consisting of a gradual change in the exemplar cluster (Beckner and Bybee 2009; Haspelmath 1998).

Thus the study of grammaticalization provides the explanatory basis for grammar as an emergent phenomenon; it also provides us with an understanding of the semantic categories of grammar and how they evolve, and an explanation for the correspondence between behavioral properties of grammatical elements and their meanings or functions (Bybee et al. 1994).

32.8 Language acquisition and adult-based change

As described in the previous sections, usage-based theory is fundamentally concerned with diachronic change, insofar as language use shapes language structure in an ongoing and dynamic fashion. A usage-based model assigns a central role to usage by *adult speakers* in accounting for language change, in contrast with the traditional generative approach, in which language change is introduced via acquisition across generations, as learners deduce a new grammar on the basis of adult speech (see Halle 1962; Lightfoot 2006, among others). In this section we note the weaknesses in the theory that allows change to occur only in the acquisition process and note the many arguments in favor of the proposal that adults can also change language.

First, a model in which children innovate via imperfect learning is unable to account for known diachronic regularities. As shown in section 32.5.1, high-frequency words and word sequences undergo the greatest degree of phonetic reduction. Such a pattern is fundamentally at odds with an imperfect learning mechanism, which predicts that children will be more likely to change low-frequency items. As we have noted, the reduction of frequent items in fact arises out of expert fluency, when well-practiced articulatory routines lead to the diminishment and overlap of speech gestures. Such reductive changes may actually result in forms that are articulatorily more complex, and harder to acquire, than the non-reduced forms. For example, consider English contracted auxiliary/negation sequences such as *did not* > *didn't* and *could not* > *couldn't*. This evidence does not point toward young language learners as the originators of the change but rather indicates that reductive change originates in usage-based factors.

Additionally, small children are unlikely to be the instigators of changes involving domains that are cognitively accessible only to older speakers. For instance, complementation is not acquired early by children, and young speakers strongly disfavor the use of overt complementizers (Radford 1990; Adamson 1992; Penke 2001). Historically, complementizers originate from a variety of lexical sources, including demonstratives (e.g., English *that*), dative-allative particles, and the verb meaning "say" (Heine and Kuteva 2002; Hopper and Traugott 2003). It seems unlikely that first language learners would misapprehend these forms as complementizers, given that young children in fact struggle to master complementation as a feature of adult language. Similarly, small children are unlikely to contribute to the grammaticalization of epistemic markers from deontics, since epistemicity cannot be fully acquired until children develop basic social competencies, in addition to attaining cognitive milestones such as a theory of mind (Barbieri and Bascelli 2000; Aksu-Koç and Alici 2000; Resches and Pereira 2007). In general, adults have more sophisticated social and cognitive abilities than children, and they face a broader range of domains in which they must communicate. Adults bring to this task a full set of capacities that can influence language change via usage, including the ability to invite and comprehend conversational inferences (Traugott 1989; see section 32.7.1 and 32.7.3). It is not plausible to assume that adults are stuck with using only the grammatical structures and conventions that children have managed to innovate. More reasonably, adult speakers are capable of extending existing patterns, which with repetition may then lead to new grammatical conventions.

Finally, there is the mismatch between children's innovation and documented diachronic changes that has been often noted in the past. Children often produce words with consonant harmony, while adult languages never have such a process (Drachman 1978; Vihman 1980); children's morphological formations at times reflect possible historical changes, but at times do not (Bybee and Slobin 1982). The fact is that children's innovations typically do not influence language because there is no social mechanism for the propagation of these innovations, given that children copy adults rather than the other way around.

The usage-based model, as described in section 32.3, proposes that even in adulthood our experiences with language continue to affect mental representations, just as in other experiential domains. It is indeed the case that adults are less influenced by new input than children, due to the cumulative effect of past tokens of experience. Moreover, early exposure to language, and early exposure to particular language features, affords learners the greatest opportunity to process language fluently (Morford 2003). However, this is not the same as saying that adult grammar is "frozen" beyond some critical period cutoff. Rather, there is a "sensitive period" that leads to a gradual decline in receptivity, but the system nonetheless never becomes completely static (Newport 1991; Morford 2002).

Contrary to claims that adults cannot adjust their grammar (Newmeyer 1998; Lightfoot 2006), we now have considerable evidence that adults continue to learn across all domains of language. A number of studies have found that speakers can adopt ongoing phonetic, even phonemic, changes in their language, long after the speaker enters adulthood (Harrington et al. 2000; Sankoff 2004; Harrington 2006; Sankoff and Blondeau 2007). Moreover, adults are not just capable of generalizing constructions to new items but must do so to use language productively. Subjects in an experiment by Kaschak and Glenburg (2004) learned an unfamiliar construction (*The meal needs cooked*) and quickly generalized this construction to new verbs. Likewise, Goldberg et al. (2004) found that, with three minutes of training, adult English speakers were able to learn an SOV construction and extend its semantics to new verbs.

In sum, we find that adult speakers are capable of participating in language change, and in some cases, adult speakers *must* be the originators of change. However, this is not to say that acquisition plays no role at all in diachronic processes. Changes such as the regularization or loss of infrequent forms may plausibly be influenced by usage (due to speakers' inability to retrieve weakly-represented variants) and by acquisition (due to children's insufficient exposure to rare variants) (Bybee and Slobin 1982). As we discuss below, a usage-based model considers contributions from multiple interacting factors in an emergentist account of language, and our catalog of language change mechanisms should be inclusive where appropriate.

32.9 Language as a complex adaptive system

In the usage-based framework, properties of languages and their grammars are viewed as emergent, i.e., not given a priori, but coming about through language use and the way the brain responds to the experience of language use (Hopper 1987; Lindblom et al. 1984; Larsen-Freeman 1997; Ellis and Larsen-Freeman 2006). Emergence is a feature of complex adaptive systems—systems in which a few causal mechanisms interact iteratively to produce what appears to be structure (Holland 1998; Camazine et al. 2001). Waves on water and dunes of sand are examples: while we perceive the structure in the waves or the dunes of sand, we know that it is not given a priori that waves or dunes should have a certain structure but rather a result of the physical properties of water and sand interacting iteratively over time and space with the bottom of the sea, the wind, and so on. It might also be noted that waves and dunes show much variability and gradience and, while we can recognize them when we see them, it might be difficult to give a firm description of their apparent structure.

We have tried to make the case in this chapter that what we perceive as language structure comes about through the application of a handful of common mechanisms that recur when human beings use language. The domain-general processes of sequential learning, chunking, categorization, and inference-making, along with the effect of partial or complete repetition, lead to the establishment and conventionalization of the categories and structures we find in languages. This bottom-up and emergentist perspective, we argue, may turn out to be indispensable to our understanding of linguistic processes and structure. Here it is helpful to draw a parallel with what is perhaps the best-studied complex adaptive system, namely, biological evolution. In the oft-cited slogan of Theodosius Dobzhansky (1973), "Nothing in biology makes sense except in the light of evolution". To truly understand the modern-day diversity of biological species, it is essential to take note of a range of simple interactions that contribute to causal mechanisms such as natural selection, sexual selection, and genetic drift. In the domain of language, the mechanisms of change are quite different, but in describing linguistic phenomena we must likewise take account of the interaction of simple elements, along with considering diachronic processes as a source of explanations. Ignoring such considerations and defaulting to a nativist, top-down explanation runs the risk that we will overlook important regularities that emerge from diachrony.

As Greenberg has argued (1969; 1978*a*; 1978*b*) the source of structure in phonology and grammar and the explanation for their similarities across languages is the set of diachronic processes that are common cross-linguistically. Commonly-occurring sound changes create phonemic systems and the cross-linguistic markedness patterns they exhibit (Greenberg 1969; and for a more work in this tradition, Bybee 2001*a*; Blevins 2004). The major source of grammatical structure is the set of processes that constitute grammaticalization. A striking characteristic of grammaticalization is that very similar developments take place in different, unrelated languages (Bybee et al. 1994; Heine and Kuteva 2002). For instance, Bybee et al. (1994) found instances of a future marker developed from a movement verb in seventeen languages out of seventy-six languages chosen to be maximally unrelated. So the development of English *be going to* into a future is not an isolated occurrence but rather reflects a very strong tendency across languages. It is possible with these and similar results to construct diachronic paths of change that are cross-linguistically similar though not always identical, perhaps in the way that one sand dune resembles another without being identical to it. An historical approach holds the most promise for explaining the complex patterns we find across the world's languages: absolute synchronic universals are rather rare (Croft 2003; Bybee 2008), but there is indeed patterned variation, in the form of statistical tendencies and recurrent diachronic shifts.

Specific unidirectional paths of change for the grammaticalization of tense, aspect, modality, voice, definites and indefinites, and many other categories have been hypothesized based on both diachronic and cross-linguistic data. While much has been written about such paths of change (Greenberg 1978*a*; Givón 1979; Bybee et al. 1994; Heine and Kuteva 2002), in our theoretical perspective, they are not at all the end of the story. More important are the mechanisms that create these paths and they are precisely the domain-general processes we mentioned earlier—chunking, categorization, inference-making, generalization (Bybee 2006*b*). So when it comes to understanding how languages are alike and how they are different, it is important once again to take a diachronic perspective and to see how the processes that create the units and structures of language interact to give us the full range of types of human language.

Grammaticalization paths and the paths of phonological change (i.e., sound change) can be thought of as substantive universals, as they refer directly to linguistic substance of phonetic form and meaning. There are also universal paths that create structural or formal universals, such as Structure Preservation (Kiparsky 1985). While this does not seem to be an absolute universal, it does express a strong tendency, which is that segments involved in lexical or morphological alternations are phonemes in their own right. Bybee (2008) demonstrates that this tendency is a result of parallel developments along several paths of change, including the unidirectional tendency for phonetic changes to become associated with morphology and lexicon. Again, the paths of change themselves are the result of the application of a handful of mechanisms that operate as language is used in context.

Usage-based theory and a complex adaptive systems approach also allows us to find explanations for correspondences that are not incorporated into other theories. For instance, the observations that grammatical morphemes (function words and affixes) are usually short (comprised of fewer segments than lexical items in the same language) and highly frequent are both observations that have a direct explanation when the usage factors in grammaticalization are taken into account (see section 32.7.3). Grammaticalization does not take place without extreme frequency increases; these same frequency increases lead to phonetic reduction.

Thus usage-based theory views language as fluid and dynamic, changing through the interaction of social usage events with the cognitive processes characteristic of the human brain in general. We have tried to show here how fundamental cognitive processes apply to linguistic experience to create the range of units and categories exhibited in human language, those structural properties that have intrigued linguists for centuries. The basic elements we have sketched, incorporating gradient categories, exemplar storage, and a non-static representational system, can account for the striking dual nature of language, as a system that undergoes change in systematic ways, while also exhibiting sufficient stability to allow communication.

This page intentionally left blank

CHAPTER 33

WORD GRAMMAR

RICHARD HUDSON

33.1 A brief history of Word Grammar (WG)

Among the questions that we have been asked to consider is question (n): "How does your model relate to alternative models?" Very few of the ideas in Word Grammar (WG) are original so it may be helpful to introduce the theory via the various theories from which the main ideas come.

We start with the name "Word Grammar" (WG), which is less informative now than it was in the early 1980s when I first used it (Hudson 1984). At that time, WG was primarily a theory of grammar in which words played a particularly important role (as the only units of syntax and the largest of morphology). At that time I had just learned about dependency grammar (Anderson 1971; Ágel and Fischer this volume), which gave me the idea that syntax is built around words rather than phrases (see section 33.8). But the earlier roots of WG lie in a theory that I had called "Daughter-Dependency Grammar" (Hudson 1976, Schachter 1978; Schachter 1981) in recognition of the combined roles of dependency and the "daughter" relations of phrase structure. This had in turn derived from the first theory that I learned and used, Systemic Grammar (which later turned into Systemic Functional Grammar—Halliday 1961; Hudson 1971; Caffarel this volume). Another WG idea that I derived

I should like to thank Nik Gisborne for help with this chapter. Interested readers will find a great deal more information on the Word Grammar website at www.phon.ucl.ac.uk/home/dick/wg.htm, and many of the papers I refer to can be downloaded from www.phon.ucl.ac.uk/home/dick/papers.htm.

from Systemic Grammar is that "realization" is different from "part", though this distinction is also part of the more general European tradition embodied in the "word-and-paradigm" model of morphology (Robins 2001; Hudson 1973).

In several respects, therefore, early WG was a typical "European" theory of language based on dependency relations in syntax and realization relations in morphology. However, it also incorporated two important American innovations. One was the idea that a grammar could, and should, be generative (in the sense of a fully explicit grammar that can "generate" well-formed structures). This idea came (of course) from what was then called Transformational Grammar (Chomsky 1965), and my first book was also the first of a series of attempts to build generative versions of Systemic Grammar (Hudson 1971). This concern for theoretical and structural consistency and explicitness is still important in WG, as I explain in section 33.2. The second American import into WG is probably its most general and important idea: that language is a network (Hudson 1984: 1; 2007*b*: 1). Although the idea was already implicit in the "system networks" of Systemic Grammar, the main inspiration was Stratificational Grammar (Lamb 1966). I develop this idea in section 33.3.

By 1984, then, WG already incorporated four ideas about grammar in a fairly narrow sense: two European ideas (syntactic dependency and realization) and two American ones (generativity and networks). But even in 1984 the theory looked beyond grammar. Like most other contemporary theories of language structure, it included a serious concern for semantics as a separate level of analysis from syntax; so in Hudson (1984), the chapter on semantics has about the same length as the one on syntax. But more controversially, it rejected the claim that language is a unique mental organ in favor of the (to my mind) much more interesting claim that language shares the properties of other kinds of cognition (Hudson 1984: 36, where I refer to Lakoff 1977). One example of a shared property is the logic of classification, which I then described in terms of "models" and their "instances", which "inherit" from the models (Hudson 1984: 14–21) in a way that allows exceptions and produces "prototype effects" (ibid. 39–41). These ideas came from my elementary reading in artificial intelligence and cognitive science (e.g., Winograd 1972; Quillian and Collins 1969; Schank and Abelson 1977); but nowadays I describe them in terms of the "isa" relation of cognitive science (Reisberg 2007) interpreted by the logic of multiple default inheritance (Luger and Stubblefield 1993: 387); section 33.4 expands these ideas.

The theory has developed in various ways since the 1980s. Apart from refinements in the elements mentioned above, it has been heavily influenced by the "cognitive linguistics" movement (Geeraerts and Cuyckens 2007; Bybee and Beckner, Fillmore, and Langacker this volume). This influence has affected the WG theories of lexical semantics (section 33.9) and of learning (section 33.10), both of which presuppose that language structure is deeply embedded in other kinds of cognitive structures. Another development has been in the theory of processing, where I have tried to take account of elementary psycholinguistics (Harley 1995), as I explain in section 33.10. But perhaps the most surprising source of influence has been sociolinguistics, in which I have a long-standing interest (Hudson 1980; 1996). I describe this influence as surprising because sociolinguistics has otherwise had virtually no impact on theories of language structure. WG, in contrast, has always been able to provide a theoretically motivated place for sociolinguistically important properties of words such as their speaker and their time (Hudson 1984: 242; 1990: 63–6; 2007*b*: 236–48). I discuss sociolinguistics in section 33.11.

In short, WG has evolved over nearly three decades by borrowing ideas not only from a selection of other theories of language structure ranging from Systemic Functional Grammar to Generative Grammar but also from artificial intelligence, psycholinguistics, and sociolinguistics. I hope the result is not simply a mishmash of ideas but an integrated framework of ideas. On the negative side, the theory has research gaps including phonology, language change, metaphor, and typology. I hope others will be able to fill these gaps. However, I suspect the main gap is a methodological one: the lack of suitable computer software for holding and testing the complex systems that emerge from serious descriptive work.

33.2 The aims of analysis

This section addresses the following questions that the editors of this volume presented in chapter 1:

(a) How can the main goals of your model be summarized?
(b) What are the central questions that linguistic science should pursue in the study of language?
(e) How is the interaction between cognition and grammar defined?
(f) What counts as evidence in your model?
(m) What kind of explanations does your model offer?

Each of the answers will revolve around the same notion: psychological reality.

Starting with question (a), the main goal of WG, as for many of the other theories described in this book, is to explain the structure of language. It asks what the elements of language are and how they are related to one another. One of the difficulties in answering these questions is that language is very complicated, but another is that we all have a number of different, and conflicting, mental models of language, including the models that Chomsky has called "E-language" and "I-language" (Chomsky 1986*a*). For example, if I learn (say) Portuguese from a book, what I learn is a set of words, rules, and so on which someone has codified as abstractions; in that case, it makes no sense to ask "Where is Portuguese?" or "Who does Portuguese belong to?" There is a long tradition of studying languages—especially dead languages—in precisely this way, and the tradition lives on in modern linguistics whenever we describe "a language". This is "external" E-language, in contrast with the purely internal I-language of a given individual, the knowledge which they hold in their brain. As with most other linguistic theories (but not Systemic Functional Grammar), it is I-language rather than E-language that WG tries to explain.

This goal raises serious questions about evidence—question (f)—because in principle, each individual has a unique language, though since we learn our language from other people, individual languages tend to be so similar that we can often assume that they are identical. If each speaker has a unique I-language, evidence from one speaker is strictly speaking irrelevant to any other speaker; and, in fact, any detailed analysis is guaranteed eventually to reveal unsuspected differences between speakers. On the other hand, there are close limits to this variation set by the fact that speakers try extraordinarily hard to conform to their role-models (Hudson 1996: 10–14), and we now know, thanks to sociolinguistics, a great deal about the kinds of similarities and differences that are to be expected among individuals in a community. This being so, it is a fair assumption that any expert speaker (i.e., barring children and new arrivals) speaks for the whole community until there is evidence to the contrary. The assumption may be wrong in particular cases, but without it descriptive linguistics would grind to a halt. Moreover, taking individuals as representative speakers fits the cognitive assumptions of theories such as WG because it allows us also to take account of experimental and behavioral evidence from individual subjects. This is important if we want to decide, for example, whether regular forms are stored or computed (Bybee 1995)—a question that makes no sense in terms of E-language. In contrast, it is much harder to use corpus data as evidence for I-language because it is so far removed from individual speakers or writers.

As far as the central questions for linguistic science—question (b)—are concerned, therefore, they all revolve around the structure of cognition. How is the "language" area of cognition structured? Why is it structured as it is? How does this area relate to other areas? How do we learn it, and how do we use it in speaking and listening (and writing and reading)? This is pure science, the pursuit of understanding for its own sake, but it clearly has important consequences for all sorts of practical activities. In education, for instance, how does language grow through the school years, and how does (or should) teaching affect this growth? In speech and language therapy, how do structural problems cause problems in speaking and listening, and what can be done about them? In natural-language processing by computer, what structures and processes would be needed in a system that worked just like a human mind?

What, then, of the interaction between cognition and grammar—question (e)? If grammar is part of cognition, the question should perhaps be: How does grammar interact with the rest of cognition? According to WG, there are two kinds of interaction. On the one hand, grammar makes use of the same formal cognitive apparatus as the rest of cognition, such as the logic of default inheritance (section 33.4), so nothing prevents grammar from being linked directly to other cognitive areas. Most obviously, individual grammatical constructions may be linked to particular types of context (e.g., formal or informal) and even to the conceptual counterparts of particular emotions (e.g., the construction *WH X*, as in *What on earth are you doing?*, where X must express an emotion; cf. Kay and Fillmore 1999 on the *What's X doing Y* construction). On the other hand, the intimate connection between grammar and the rest of cognition allows grammar to influence non-linguistic cognitive development as predicted by the Sapir–Whorf hypothesis (Lee 1996; Levinson 1996*b*). One possible consequence of this influence is a special area of cognition outside language which is only used when we process language—Slobin's "thinking for speaking" (Slobin 1996). More generally, a network model predicts that some parts of cognition are "nearer" to language (i.e., more directly related to it) than others, and that the nearer language is, the more influence it has.

Finally, we have the question of explanations—question (m). The best way to explain some phenomenon is to show that it is a special case of some more general phenomenon, from which it inherits all its properties. This is why I find nativist explanations in terms of a unique "language module" deeply unsatisfying, in contrast with the research program of cognitive linguistics whose basic premise is that "knowledge of language is knowledge" (Goldberg 1995: 5). If this premise is true, then we should be able to explain all the characteristics of language either as characteristics shared by all knowledge, or as the result of structural pressures from the ways in which we learn and use language. So far I believe the results of this research program are very promising.

33.3 Categories in a Network

As already mentioned in section 33.1, the most general claim of WG is that language is a network, and more generally still, knowledge is a network. It is important to be clear about this claim because it may sound harmlessly similar to the structuralist idea that language is a system of interconnected units, which every linguist would accept. It is probably uncontroversial that vocabulary items are related in a network of phonological, syntactic, and semantic links, and networks play an important part in the grammatical structures of several other theories (notably system networks in Systemic Functional Grammar and directed acyclic graphs in Head-Driven Phrase Structure Grammar—Pollard and Sag 1994). In contrast with these theories where networks play just a limited part, WG makes a much bolder claim: in language there is nothing but a network—no rules or principles or parameters or processes, except those that are expressed in terms of the network. Moreover, it is not just the language itself that is a network; the same is true of sentence structure, and indeed the structure of a sentence is a temporary part of the permanent network of the language. As far as I know, the only other theory which shares the view that "it's networks all the way down" is Neurocognitive Linguistics (Lamb 1998).

Moreover, the nodes of a WG network are atoms without any internal structure, so a language is not a network of complex information-packages such as lexical entries or constructions or schemas or signs. Instead, the information in each such package must be "unpacked" so that it can be integrated into the general network. The difference may seem small, involving little more than the metaphor we choose for talking about structures; but it makes a great difference to the theory. If internally complex nodes are permitted, then we need to allow for them in the theory by providing a typology of nodes and node-structures, and mechanisms for learning and exploiting these node-internal structures. But if nodes are atomic, there is some hope of providing a unified theory which applies to all structures and all nodes.

To make the discussion more concrete, consider the network-fragment containing the synonyms BEAR$_{verb}$ and TOLERATE and the homonyms BEAR$_{verb}$ and BEAR$_{noun}$ (as in *I can't bear the pain* and *The bear ate the honey*). The analysis in Figure 33.1 is in the spirit of Cognitive Grammar (e.g., Langacker 1998*b*: 16), so it recognizes three "symbolic units" with an internal structure consisting of a meaning (in quotation marks) and a form (in curly brackets). Since symbolic units cannot overlap, the only way to relate these units to each other is to invoke separate links to other units in which the meanings and forms are specified on their own. In this case, the theory must distinguish the relations between units from those found within units, and must say what kinds of units (apart from symbolic units) are possible.

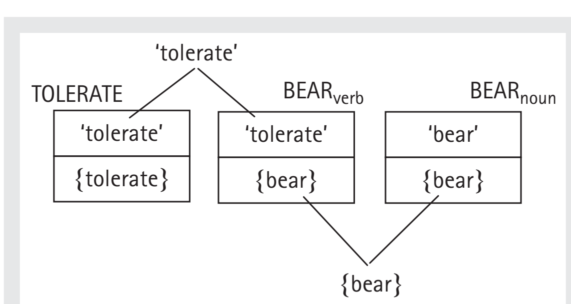


Figure 33.1. Two synonyms and two homonyms as a network of complex units

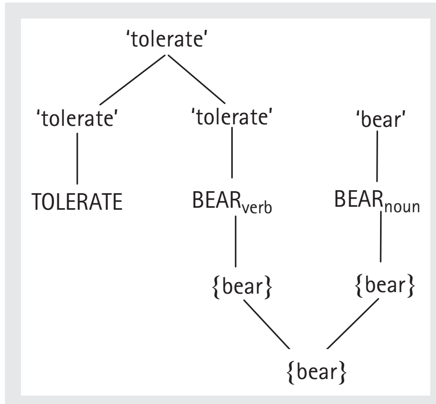


Figure 33.2. Two synonyms and two homonyms as a pure network

This analysis can be contrasted with the one in Figure 33.2, which is in the spirit of WG but does not use WG notation (for which see Figure 33.3 below). In this diagram there are no boxes because there are no complex units—just atomic linked nodes. The analysis still distinguishes different kinds of relations and elements, but does not do it in terms of boxes. The result is a very much simpler theory of cognitive structure in which the familiar complexes of language such as lexical items and constructions can be defined in terms of atomic units.

We can now turn to question (c): "What kinds of categories are distinguished?" WG recognises three basic kinds of elements in a network:

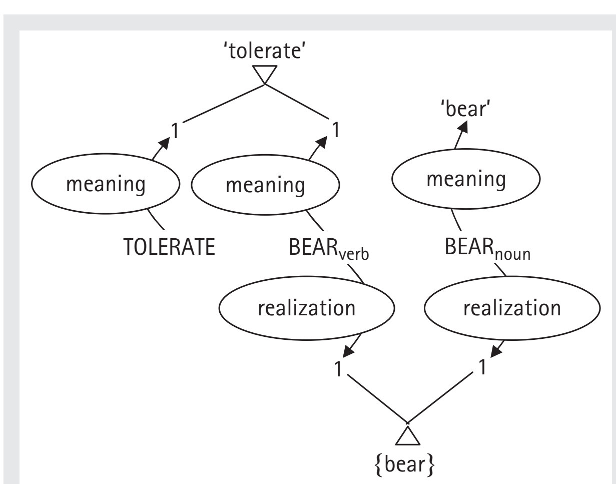


Figure 33.3. Two synonyms and two homonyms in WG notation

- Primitive logical relations: "isa" (the basic relation of classification which Langacker calls "schematicity"; (Tuggy 2007)) and four others: "identity", "argument", "value", and "quantity" (Hudson 2006: 47).
- Relational concepts: all other relations whether linguistic (e.g., "meaning", "realization", "complement") or not (e.g., "end", "father", "owner").
- Non-relational concepts, whether linguistic (e.g., "noun", "{bear}", "singular") or not (e.g., "bear", "tolerate", "set").

The "isa" relation plays a special role because every concept, whether relational or not, is part of an "isa hierarchy" which relates it upward to more general concepts and downward to more specific concepts. For example, "complement" isa "dependent", and "object" isa "complement", so the network includes a hierarchy with "complement" above "object" and below "dependent". As I explain in section 33.4, "isa" also carries the basic logic of generalization, default inheritance.

Any network analysis needs a notation which distinguishes these basic types of element. The WG notation which does this can be seen in Figure 33.3:

- Relational concepts are named inside an ellipse.
- Non-relational concepts have labels with no ellipse.
- Primitive logical relations have distinct types of line. The "isa" relation has a small triangle whose base rests on the super-category; "argument" and "value" are the arcs pointing into and out of the relational concept; and "quantity" is shown (without any line) by a digit which represents a non-relational concept.

In other words, therefore, the figure shows that the meaning of the noun BEAR ($\text{BEAR}_{\text{noun}}$) is "bear"; and because "tolerate" may be the meaning of either TOLERATE or the verb BEAR, two different instance of "tolerate" are distinguished so that each is the meaning of a different verb. This apparently pointless complexity is required by the logic of WG, which otherwise cannot express the logical relation "or"—see section 33.4.

33.4 The logic of inheritance

As in any other theory, the linguist's analysis tries to capture generalizations across words and sentences in the language concerned, so the mechanism for generalization plays a crucial role. Since the goal of the analysis is psychological reality in linguistic analysis combined with the attempt to use general-purpose cognitive machinery wherever possible, the mechanism assumed in WG is that of everyday reasoning, and default inheritance (Pelletier and Elio 2005). The same general principle is assumed in a number of other linguistic theories (Pollard and Sag 1994: 36; Jackendoff 2002: 184; Goldberg 2006: 171; Bouma 2006).

The general idea is obvious and probably uncontroversial when applied to common-sense examples. For example, a famous experiment found that people were willing to say that a robin has skin and a heart even though they did not know this as a fact about robins as such. What they did know, of course, was, first, that robins are birds and birds are living creatures ("animals" in the most general sense), and, second, that the typical animal (in this sense) has skin and a heart (Quillian and Collins 1969). In other words, the subjects had "inherited" information from a super-category onto the sub-category. We all engage in this kind of reasoning every minute of our lives, but we know that there are exceptions which may prove us wrong—and, indeed, it is the exceptions that make life both dangerous and interesting. If inheritance allows for exceptions, then it is called "default inheritance" because it only inherits properties "by default", in the absence of any more specific information to the contrary. This is the kind of logic that we apply in dealing with familiar "prototype effects" in categorization (Rosch 1978); so if robins are more typical birds than penguins, this is because penguins have more exceptional characteristics than robins do. Somewhat more precisely, the logic that we use in everyday life allows one item to inherit from a number of super-categories; for example, a cat inherits some characteristics from "mammal" (e.g., having four legs) and others from "pet" (e.g., living indoors with humans). This extension of default inheritance is called "multiple default inheritance".

It is reasonably obvious that something like this logic is also needed for language structure, where exceptions are all too familiar in irregular morphology, in "quirky" case selection, and so on, and where multiple inheritance is commonplace—for instance, a feminine, accusative, plural noun inherits independently from "feminine", "accusative", and "plural". This logic is implied by the "Elsewhere condition" (Kiparsky 1982) in lexical phonology, and is implicit in many other approaches such as rule-ordering where later (more specific) rules can overturn earlier more general ones. Nevertheless, multiple default inheritance is considered problematic in linguistic theory, and much less widely invoked than one might expect. One reason for this situation is the difficulty of reconciling it with standard logic. Standardly, logic is "monotonic", which means that, once an inference is drawn, it can be trusted. In contrast, default inheritance is non-monotonic because an inference may turn out to be invalid because of some exception that overrides it. Moreover, multiple inheritance raises special problems when conflicting properties can be inherited from different super-categories (Touretzky 1986). WG avoids these logical problems (and others) by a simple limitation: inheritance only applies to tokens (Hudson 2006: 25). How this works is explained below.

To take a simple linguistic example, how can we show that by default the past tense of a verb consists of that verb's stem followed by the suffix {ed}, but that for TAKE the past tense form is not *taked* but *took*? The WG answer is shown in Figure 33.4. The default pattern is shown in the top right-hand section: "past" (the

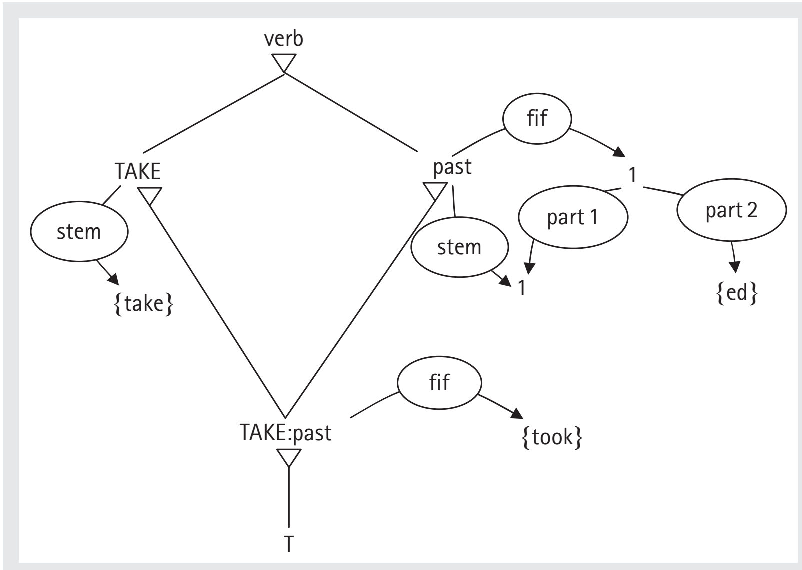


Figure 33.4. An irregular verb overrides the default past tense form

typical past tense verb) has a "fully inflected form" (fif) consisting of the verb's stem followed by {ed}. The entry for TAKE in the top left shows that its stem is {take}, so by default the fif of a word which inherits (by multiple inheritance) from both TAKE and "past" should be {{take}{ed}}. However, the fif is in fact specified as {took}, so this form overrides the default. Now suppose we apply this analysis to a particular token T which is being processed either in speaking or in listening. This is shown in the diagram with an isa link to TAKE:past, as explained in section 33.10. If inheritance applies to T, it will inherit all the properties above it in the hierarchy, including the specified fif; but the process inevitably starts at the bottom of the hierarchy so it will always find overriding exceptions before it finds the default. This being so, the logic is actually monotonic: once an inference is drawn, it can be trusted.

Default inheritance is important in linguistic analysis because it captures the asymmetrical relation which is found between so many pairs of alternatives, and which in other theories is expressed as one of the alternatives being the "underlying" or "unmarked" one. For example, one word order can be specified as the default with more specific orders overriding it; so a dependent of an English word typically follows it, but exceptionally the subject of a verb typically precedes it, but exceptionally the subject of an "inverting" auxiliary verb typically follows it (see section 33.8 for word order). The same approach works well in explaining the complex ordering of extracted words in Zapotec, as well as a wide range of other asymmetrical patterns (Hudson 2003*c*).

Another role of default inheritance is to capture universal quantification. If X has property P, then "all X", i.e., everything which isa X, also has property P. The main difference is that, unlike universal quantification, default inheritance allows exceptions. In contrast, the WG equivalent of the other kind of quantification, existential quantification, is simply separate "existence" in the network; so if "some X" has property P, there is a separate node Y in the network which isa X and has the property P. Other examples of X do not inherit P from Y because there is no "upward inheritance". Similarly, inheritance makes the "and" relation easy to express: if X has two properties P and Q, then both are automatically inherited by any instance of X. In contrast, the relation "or" is much harder to capture in a network—as one might hope, given its relative complexity and rarity. The solution in WG is to recognize a separate sub-case for each of the alternatives; so if X has either P or Q among its properties, we assign each alternative to a different sub-case of X, X1, and X2—hence the two sub-cases of {bear} in Figure 33.3.

33.5 The architecture of language

The formal structure of WG networks described in section 33.3 already implies that they have a great deal of structure because every element is classified hierarchically. This allows us to distinguish the familiar levels of language according to the vocabulary of units that they recognize: words in syntax, morphs in morphology, and phones in phonology. Moreover, different relation-types are found on and between different levels, so levels of analysis are at least as clearly distinguished in WG as they are in any other theory. This allows us to consider question (d): "What is the relation between lexicon, morphology, syntax, semantics, pragmatics, and phonology?"

We start with the lexicon. WG (just like other cognitive theories—Croft 2007*a*: 471) recognizes no boundary between lexical and "grammatical" structures; instead, it simply recognizes more and less general word-types. For example, the verb $\text{BEAR}_{\text{verb}}$ isa Transitive-verb, which isa Verb, which isa Word, and at no point do we find a qualitative difference between specific "lexical" and general "grammatical" concepts. Nor can we use length as a basis for distinguishing one-word lexical items from multi-word general constructions because we clearly memorize individual multi-word idioms, specific constructions, and clichés. Moreover, almost every theory nowadays recognizes that lexical items have a valency which defines virtual dependency links to other words, so all "the grammar" has to do is to "merge" lexical items so that these dependencies are satisfied (Ninio 2006: 6–10; Chomsky

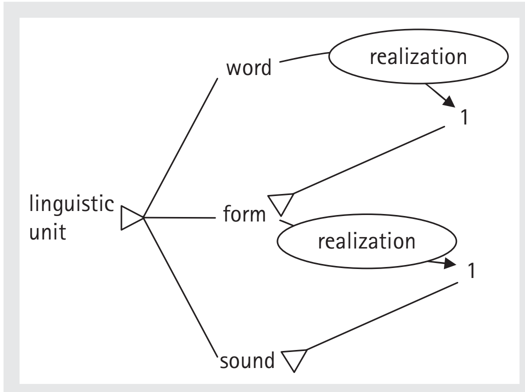


Figure 33.5. The three linguistic levels in WG notation

1995: 226)—a process that involves nothing more specific than ensuring that the properties of a token (such as its dependents) match those of its type. In short, the syntactic part of the language network is just a highly structured and hierarchical lexicon which includes relatively general entries as well as relatively specific ones (Flickinger 1987)—what we might call a "super-lexicon".

However, WG does not recognize just one super-lexicon specific to language but three: one for syntax (consisting of words), another for morphology, and a third for phonology. The morphological lexicon consists of what I call "forms"—morphs such as {bear}, {bore}, and {s}, and morph-combinations extending up to complete word-forms such as {{un}{bear}{able}} and {{walk}{s}} (Hudson 2006: 72–81). In phonology, I assume the vocabulary of units includes segments and syllables, but in WG this is unexplored territory. This analysis gives a three-level analysis within language; for example, the word FARMER:plural (the plural of FARMER) is realized by the form {{farm}{er}{s}} which in turn is realized by a phonological structure such as /fɑː/məz/. Each level is identified not only by the units that it recognizes but also by the units that realize them and those that they realize; so one of the characteristics of the typical word is that it is realized by a form, and by default inheritance this characteristic is inherited by any specific word. The overall architecture of WG in terms of levels is shown in Figure 33.5, where every word is realized by some form and every form is realized by some sound. (Not every form realizes a word by itself, nor does every sound realize a form by itself.) What units at all three levels share is the fact that they belong to some language (English, French, or whatever), so they are united as "linguistic units".

This three-level analysis of language structure is controversial, of course, though by no means unprecedented (Aronoff 1994; Sadock 1991). It conflicts with any analysis in terms of bipolar "signs" which combine words (or even meanings) directly with phonology (Pollard and Sag 1994; Chomsky 1995; Langacker 1998*b*; Jackendoff 1997; Beard 1994; Anderson 1992), as well as with neo-Bloomfieldian analyses which treat morphemes as word-parts (Halle and Marantz 1993). The WG claim is that the intermediate level of "form" is psychologically real, so it is encouraging that the most widely accepted model of speech processing makes the same assumption (Levelt et al. 1999). The claim rests on a variety of evidence (Hudson 2006: 74–8) ranging from the invisibility of phonology in syntax to the clear recognition of morphs in popular etymology. It does not follow from any basic principles of WG, so if it is true it raises research questions. Do all languages have the same three-level organization? For those languages that do have it, why have they evolved in this way?

A particularly controversial aspect of this three-level analysis is the place of meaning. The simplest assumption is that only words have meaning, so morphs have no meaning. This seems right for morphs such as the English suffix {s}, which signals two completely different inflectional categories (plural in nouns and singular in verbs); and if the form {bear} realizes either the verb or the noun, then there is little point in looking for its meaning. On the other hand, it is quite possible (and compatible with WG principles) that some morphs do have a meaning; and, indeed, there is experimental evidence for "phonaesthemes"—purely phonological patterns such as initial /gl/ in English that correlate with meanings, though rather more loosely than forms and words do (Bergen 2004). Moreover, intonational and other prosodic patterns have a meaning which contributes to the overall semantic structure, for instance by distinguishing questions from statements. It seems quite likely, therefore, that units at all levels can have a meaning. On the other hand, this is a typical property of words, in contrast with forms and sounds which typically have no meaning, so there is still some truth in the earlier WG claim that meanings are expressed only by words.

The default logic of WG (section 33.4) allows exceptions in every area, including the basic architecture of the system. We have just considered one example, morphological and phonological patterns that have meanings; and it cannot be ruled out that words might be realized in some cases directly by sounds. Another kind of exception is found between syntax and morphology, where the typical word is realized by a word-form (a particular kind of form which is "complete" as far as the rules of morphology are concerned). The exception here is provided by clitics, which are words—i.e., units of syntax—which are realized by affixes so that they have to be attached to other forms for the sake of morphological completeness; for example, the English possessive *_'s* (as in *John's hat*) is a determiner realized by a mere suffix. WG analyses are available for various complex clitic systems including French and Serbo-Croat pronouns (Camdzic and Hudson 2007; Hudson 2001; 2006: 104–15).

In short, WG analyzes a language as a combination of three super-lexicons for words, forms, and sounds (at different levels of generality). These lexicons are arranged hierarchically by default so that words have meanings and are typically realized by forms, and forms are typically realized by sounds, but exceptions exist. As for pragmatics, a great deal of so-called "pragmatic" information about context may be stored along with more purely linguistic properties (see sections 33.9 and 33.11), but a great deal more is computed during usage by the processes of understanding (section 33.10).

33.6 Words, features, and agreement

In the three-level analysis, the typical word stands between meaning and morphological form, so its properties include at least a meaning and a realization. However, it has other properties as well which we review briefly below.

Most words are classified in terms of the familiar super-categories traditionally described in terms of word classes (noun, verb, etc.), sub-classes (auxiliary verb, modal verb, etc.), and feature structures (tense, number, etc.). Many theories reduce all these kinds of classification to feature structures expressed as attribute-value matrices, so that a plural noun (for example) might have the value "plural" for the attribute "number" and the value "noun" for "part of speech" (or, in Chomskyan analysis, "+" for "noun" and "−" for "verb"). "Nearly all contemporary approaches use features and feature structures to describe and classify syntactic and morphological constructions" (Blevins 2006: 393). WG takes the opposite approach, using the isa hierarchy for all kinds of classification. We have already seen the effects of this principle in Figure 33.4, where both TAKE and "past" have an isa relation to "verb". This fundamental theoretical difference follows from the adoption of "isa" as the mechanism for classification, which in turn follows from the aim of treating language wherever possible like other areas of cognition. Even if attribute-value matrices are helpful in linguistic analysis, they are surely not relevant in most kinds of classification. For example, if we classify both apples and pears as a kind of fruit, what might be the attribute that distinguishes them? The problems are the same as those of the "componential analysis" that was tried, and abandoned, in the early days of modern semantics (Bolinger 1965).

Moreover, feature-based classification only works well for a very small part of language, where names such as "case" and "number" are already available for the attributes; we return to this minority of cases below. Distinctions such as the one between common and proper nouns or between auxiliary and full verbs have no traditional name, and for good reason: the "attribute" that contrasts them does no work in the grammar. Consequently, WG uses nothing but an isa hierarchy for

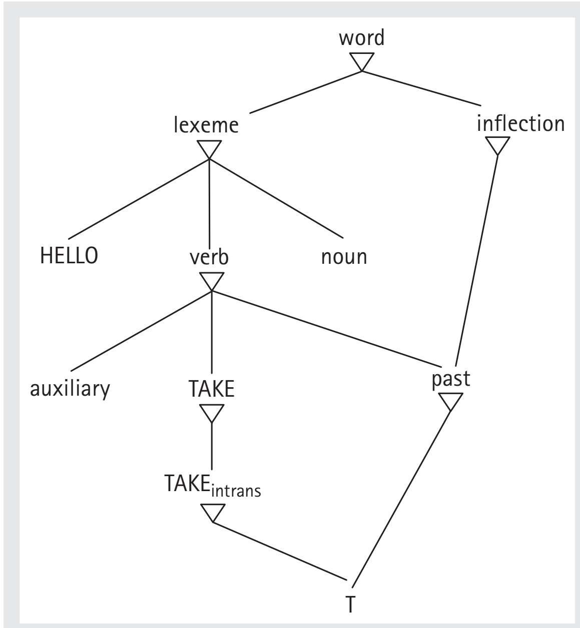


Figure 33.6. An isa hierarchy for words including classes, a sub-class, lexemes, a sub-lexeme, an inflection, and a token

classifying words. It should be borne in mind that multiple inheritance allows cross-classification, which is traditionally taken as evidence for cross-cutting attributes; for example, Figure 33.4 shows how the word TAKE:past can be classified simultaneously in terms of lexemes (TAKE) and in terms of morpho-syntactic contrasts such as tense (past). Similarly, Figure 33.6 shows how this analysis fits into a broader framework which includes:

- the super-class "word"
- very general word-types (lexeme, inflection)
- word classes (verb, noun)
- a sub-class (auxiliary)
- individual lexemes (HELLO, TAKE)
- sub-lexemes (TAKE$_{intrans}$, the intransitive use of TAKE as in *The glue wouldn't take*)
- an inflection (past)
- a word-token (T) which is analyzed as the past tense of TAKE$_{intrans}$.

This unified treatment allows the same default inheritance logic to handle all kinds of generalizations, but it also brings other advantages. First, it allows us to avoid classification altogether where there is no generalization to be captured; this is illustrated by the word HELLO, which inherits no grammatical properties from any word class, so it is "syncategorematic", belonging to no general category other than "word" (Pullum 1982). Second, default members of a category belong to that category itself, so sub-categories are only needed for exceptions. Contrary to more traditional classification systems, this means that a category may have just one sub-category. The relevant example in the diagram is "auxiliary", which does not contrast with any other word class because non-auxiliary verbs are simply default verbs. Similarly, "past" does not contrast with "present" because verbs are present tense by default; in traditional terminology, tense is a privative opposition, and "past" is marked relative to "present". Third, sub-lexemes allow distinctions without losing the unifying notion of "lexeme"; so for example it is possible to recognize both the transitive and intransitive uses of TAKE as examples of the same lexeme (with the same irregular morphology) while also recognizing the differences. And lastly, the token (which is attached temporarily to the network as explained in section 33.10) can inherit from the entire hierarchy by inheriting recursively from each of the nodes above it.

Unlike many other contemporary theories, therefore, WG classifies words without using feature-structures because, in general, they are redundant. The exception is agreement, where one word is required to have the same value as some other word for some specified attribute such as gender or number; for example, in English a determiner has the same number as its complement noun (*this book* but *these books*), and in Latin an adjective agrees with the noun on which it depends in gender, number, and case. It is impossible to express this kind of rule in a psychologically plausible way without attributes and values, but this is not a theoretical problem for WG because attributes are found in general cognition; for example, when we say that two people are the same height or age, we are invoking an attribute. Consequently, attributes are available when needed, but they are not the basis of classification—and, indeed, their relation to basic classification in the isa hierarchy may be more or less complex rather than in a simple one-to-one relation. For example, one of the values may be assigned by default, allowing the asymmetrical relations between marked and unmarked values mentioned above, which is illustrated by the default "singular" number of nouns shown in Figure 33.7. The network on the right in this figure is the English agreement rule for determiners and their complement nouns. Other agreement rules may be more complex; for example, I have suggested elsewhere that subject–verb agreement in English involves three different attributes: number, agreement-number, and subject-number, which all agree by default but which allow exceptions such as the plural verb forms used with the pronouns *I* and *you* (Hudson 1999).

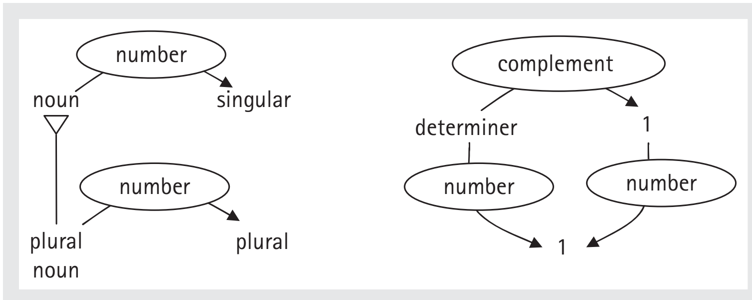


Figure 33.7. Nouns are singular by default, and a determiner agrees in number with its complement

33.7 Morphology

The three-level architecture explained in section 33.5 means that each word has a morphological structure defined in terms of morphs; this applies even to monomorphs such as CAT, realized by {cat}, which in turn is realized by /kat/. The task of morphology is to define possible morphological structures and to relate them upward to words and word classes (morpho-syntax) and downward to phonology (morpho-phonology).

In morpho-syntax, WG allows morphs to realize semantic and syntactic contrasts, but does not require this; so morphs may be purely formal objects such as the semantically opaque roots in DECEIVE and RECEIVE, where {ceive} is motivated only by the derived nouns DECEPTION and RECEPTION. In most cases, however, a word's morphological structure indicates its relations to other words with partially similar structures. The distinction between lexemes and inflections (Figure 33.6) allows two logical possibilities for these relations:

- lexical ("derivational") morphology: the two words belong to different lexemes (e.g., FARM—FARMER).
- inflectional morphology: they belong to the same lexeme (e.g., *farm—farms*).

In both cases, the partial morphological similarities may match similarities found between other lexemes.

Lexical morphology often builds on general lexical relations which exist independently of morphological structure; for example, many animal names have contrasting adult-young pairs without any morphological support (e.g., COW—CALF, SHEEP—LAMB), though in some cases the morphology is transparent (DUCK—DUCKLING, GOOSE—GOSLING). Where lexical morphology is productive, it must involve two relations: a semantically and syntactically specified lexical relation

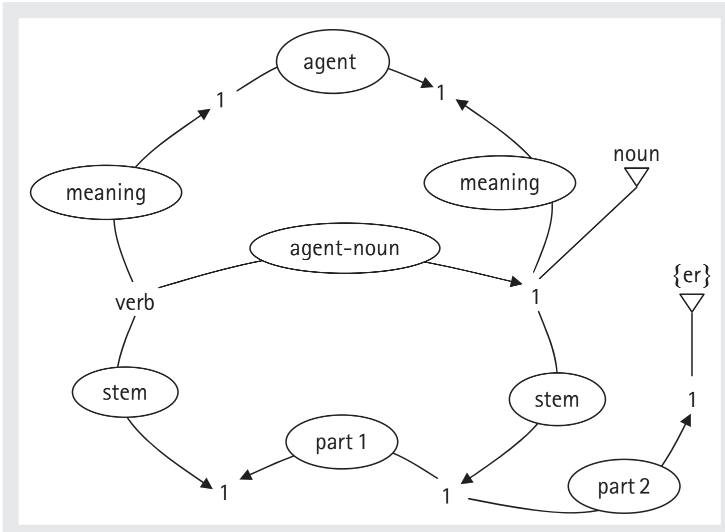


Figure 33.8. Lexical morphology: A verb is related to its agent-noun in both meaning and morphology

between two sets of words, and a morphologically specified relation between their structures. A simple example can be found in Figure 33.8, which shows that a typical verb has an "agent-noun" which defines the agent of the verb's action and whose stem consists of the verb's stem followed by {er}. (A few details in this diagram have been simplified.)

Inflectional morphology, on the other hand, relates a word's morphological structure to its inflections, the abstractions such as "past" which cut across lexical differences. As explained in section 33.1, WG follows the European "Word and Paradigm" approach to inflectional morphology by separating morphological structure from inflectional categories and avoiding the term "morpheme", which tends to confuse the two. This allows all sorts of complex mappings between the two structures, including a mapping in which several inflections are realized by a single morph (as in Latin *am-o*, "I love", where the suffix {o} realizes "first-person", "singular", "present", and "indicative").

This strict separation of morpho-syntax from morpho-phonology is not limited to inflectional morphology but runs through the entire WG approach to morphology. One consequence is that although the logical contrast between lexical and inflectional morphology applies to morpho-syntax, it is irrelevant to morpho-phonology. For example, the {er} suffix which is found in agent-nouns (Figure 33.8) is also used in the comparative inflection (as in *bigger*). In morpho-phonology the issues concern morphological structure—what kinds of structure are possible, and what kinds of generalization are needed in order to link them to sounds? The analysis deals in distinctions such as that between root morphs and affixes, and has to

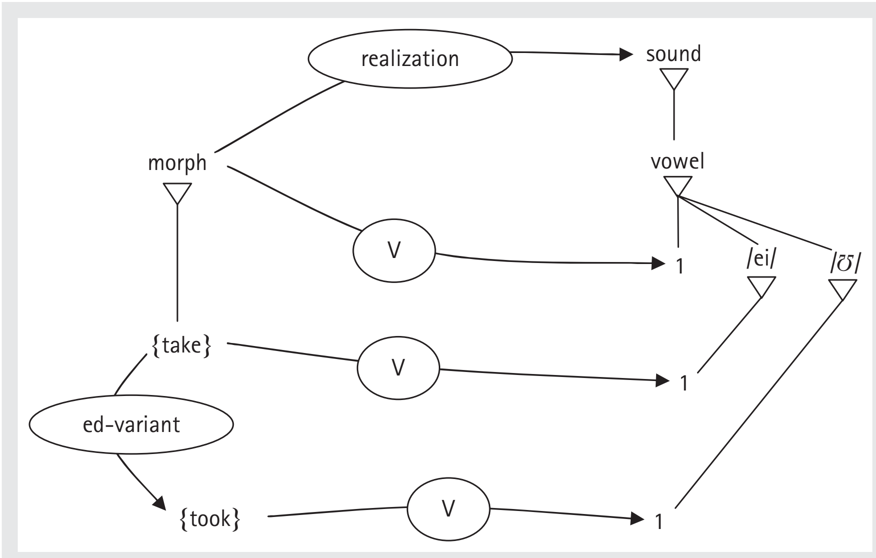


Figure 33.9. The alternation in *take—took* involves only the stressed vowel

capture generalizations such as the fact that full morphs are typically realized by one or more complete syllables, whereas affixes are often single segments. Furthermore, it has to have enough flexibility to accommodate patterns in which one structure is related to another, not by containing an extra morph but in all the other familiar ways such as vowel change as in *take—took*. We already have a partial analysis for this pair (Figure 33.4), but this simply presents {took} as an unrelated alternative to {take}, without attempting either to recognize the similarities between them or to reveal that the vowel is the usual locus for replacive morphology. Both these goals are achieved in Figure 33.9, which recognizes "V" (the stressed vowel) as a special type of realization which varies in morphs such as {take}.

This figure also illustrates another important facility in WG, the notion of a "variant". This is the WG mechanism for capturing generalizable relations between morphological structures such as that between a form and its "ed-variant"—the structure which typically contains {ed} but which may exceptionally have other forms such as the one found in {took}. Typically, a form's variant is a modification of the basic form, but in suppletion the basic form is replaced entirely by a different one. Variants have a number of uses in morpho-phonology. One is in building complex morphological structures step-wise, as when the future tense in Romance languages is said to be built on the infinitive (e.g., in French, *port-er-ai* "I will carry" but *part-ir-ai* "I will depart"). Another is in dealing with syncretism, where two or more distinct inflections systematically share the same realization; for example, in Slovene, dual and plural and plural nouns are generally different in morphology, but exceptionally the genitive and locative are always the same, and this is true even in the most irregular suppletive paradigms (Evans et al. 2001). The question is how to explain the regularity of this irregularity. One popular solution is to use a "rule of referral" (Stump 1993) which treats one form as basic and derives the other from it; so in the Slovene example, if we treat the genitive plural as basic we might use this in a rule to predict the genitive dual and locative dual. But rules of referral are very hard to take seriously if the aim is psychological reality because they imply that when we understand one form we must first mis-analyze it as a different one; and in any case, the choice of a basic form is psychologically arbitrary. The WG solution is to separate the morpho-syntax from the morpho-phonology. In morpho-phonology, we recognize a single "variant" which acts as the realization for a number of different inflections; so, for example, in Slovene, the variant which we might call (arbitrarily) "p3", and which has different morpho-phonological forms in different lexemes, is always the one used to realize dual as well as plural in the genitive and locative (Hudson 2006: 86).

The main tools in WG morphology are all abstract relations: lexical relations between lexemes, realization relations, and "variant" relations among formal structures. This is typical of a network analysis, and anticipates what we shall find in syntax.

33.8 Syntax

Syntax is the area of analysis where most work has been published in WG, and the one on which the theory's name is based (as explained in section 33.1). By far the most controversial aspect of WG syntax is the use of dependency structure instead of the more familiar phrase structure. The reason for this departure from the mainstream is that the arguments for dependency structure are very strong—in fact, even adherents of phrase structure often present it as a tool for showing syntactic dependencies—and (contrary to what I once believed—Hudson 1976) once dependencies are recognized, there are no compelling reasons for recognizing phrases as well. In WG syntax, therefore, dependencies such as "subject" or "complement" are explicit and basic, whereas phrases are merely implicit in the dependency structure. This means, for example, that the subject of a verb is always a noun, rather than a noun phrase, and that a sentence can never have a "verb phrase" (in any of the various meanings of this term). The structure in Figure 33.10 is typical of dependency relations in WG, though it does not of course try to show how the words are classified or how the whole structure is related to the underlying grammar.

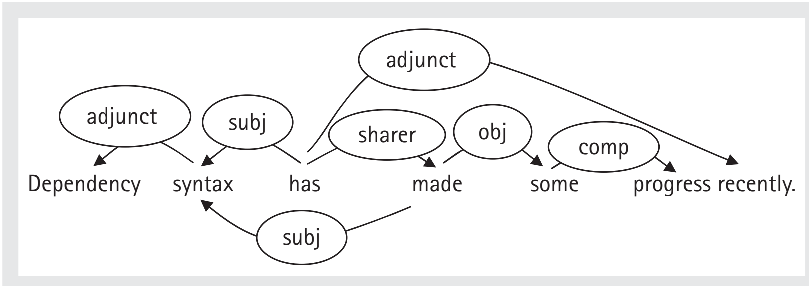


Figure 33.10. Dependency structure in an English sentence

WG dependency structures are much richer than those in other dependency grammars because their role is to reveal the sentence's entire syntactic structure rather than just one part of it (say, just semantics or just word order); and in consequence each sentence has just one syntactic structure rather than the multi-layered structures found, for example, in Functional Generative Description (Sgall et al. 1986) or the Meaning-Text Model (Mel'cuk 1997). This richness can be seen in Figure 33.10 where the word *syntax* is the subject of two verbs at the same time: *has* and *made*. The justification for this "structure sharing" (where two "structures" share the same word) is the same as in other modern theories of syntax such as Head-Driven Phrase Structure Grammar (Pollard and Sag 1994: 2). However, some WG structures are impossible to translate into any alternative theory because they involve mutual dependency—two words each of which depends on the other. The clearest example of this is in wh-questions, where the verb depends (as complement) on the wh-word, while the wh-word depends (e.g., as subject) on the verb (Hudson 2003*d*), as in Figure 33.11. Such complex structures mean that a syntactic sentence structure is a network rather than a mere tree-structure, but this is hardly surprising given that the grammar itself is a network.

Word order is handled in current WG by means of a separate structure of "landmarks" which are predicted from the dependency structure. The notion of

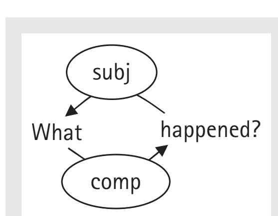


Figure 33.11. Mutual dependency in a wh-question

"landmark" is imported from Cognitive Grammar (e.g., Langacker 1990, 6), where it is applied to the semantics of spatial relations; for example, if X is in Y, then Y is the landmark for X. In WG it is generalized to syntax as well as semantics because in a syntactic structure each word takes its position from one or more other words, which therefore act as its "landmark". In the WG analysis, "before" and "after" are sub-cases of the more general "landmark" relation. By default, a word's landmark is the word it depends on, but exceptions are allowed because landmark relations are distinct from dependency relations. In particular, if a word depends on two other words, its landmark is the "higher" of them (in the obvious sense in which a word is "lower" than the word it depends on); so in Figure 33.10 the word *syntax* depends on both *has* and *made*, but only takes the former as its landmark. This is the WG equivalent of saying that *syntax* is "raised". Similarly, the choice of order relative to the landmark (between "before" and "after") can be set by default and then overridden in the way described at the end of section 33.4.

Published WG analyses of syntax have offered solutions to many of the familiar challenges of syntax such as extraction islands and coordination (see especially Hudson 1990: 354–421) and gerunds (Hudson 2003*b*). Although most analyses concern English, there are discussions of "empty categories" (in WG terms, unrealized words) in Icelandic, Russian, and Greek (Creider and Hudson 2006; Hudson 2003*a*) and of clitics in a number of languages, especially Serbo-Croatian (Camdzic and Hudson 2007; Hudson 2001).

33.9 Semantics

When WG principles are applied to a sentence's semantics they reveal a much more complex structure than the same sentence's syntactic structure. As in Frame Semantics (Fillmore, this volume), a word's meaning needs to be defined by its "frame" of relations to a number of other concepts which in turn need to be defined in the same way, so ultimately the semantic analysis of the language is inseparable from the cognitive structures of the users. Because of space limitations, all I can do here is to offer the example in Figure 33.12 with some comments and refer interested readers to other published discussions (Hudson 1990: 123–66; Hudson 2006: 211–36; Hudson and Holmes 2000; Gisborne 2001).

The example gives the syntactic and semantic structure for the sentence *The dog hid a bone for a week*. The unlabeled syntactic dependency structure is drawn immediately above the words, and the dotted arrows link the words to relevant parts of the semantic structure; although this is greatly simplified, it still manages to illustrate some of the main achievements of WG semantics. The usual "1" labels

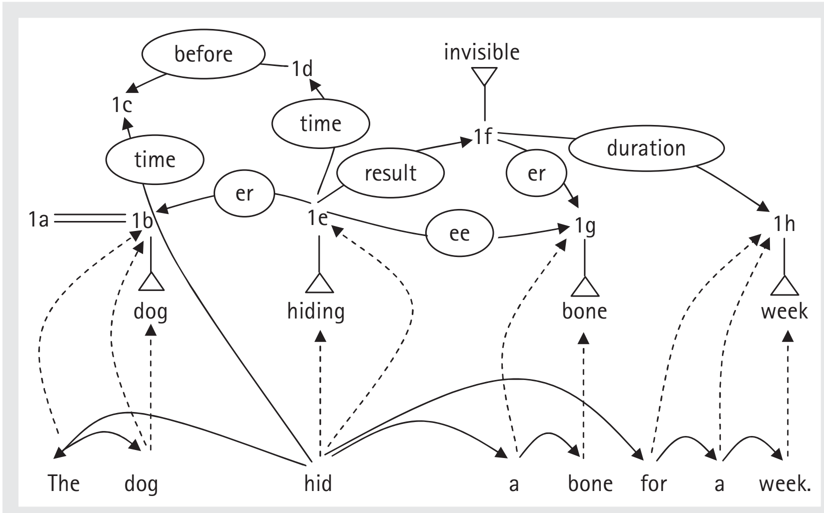


Figure 33.12. Syntactic and semantic structure for a simple English sentence

(meaning a single token) have been distinguished by a following letter for ease of reference below.

The analysis provides a mentalist version of the familiar sense/referent distinction (Jackendoff 2002: 294) in two kinds of dotted lines: straight for the sense and curved for the referent. Perhaps the most important feature of the analysis is that it allows the same treatment for all kinds of words, including verbs (whose referent is the particular incident referred to), so it allows events and other situations to have properties like those of objects; this is the WG equivalent of Davidsonian semantics (Davidson 1967; Parsons 1990). For example, "1e" shows that there was just one incident of hiding, in just the same way that "1b" shows there was just one dog.

Definiteness is shown by the long "=" line which indicates the basic relation of identity (section 33.3). This line is the main part of the semantics of *the*, and indicates that the shared referent of *the* and its complement noun needs to be identified with some existing node in the network. This is an example of WG semantics incorporating a good deal of pragmatic information. The treatment of deictic categories such as tense illustrates the same feature; in the figure "1d", the time of the boiling is before "1c", the time of the word *boiled* itself.

The decomposition of "hiding" into an action (not shown in the diagram) and a result ("invisible") solves the problem of integrating time adverbials such as *for a week* which presuppose an event with extended duration. Hiding, in itself, is a punctual event so it cannot last for a week; what has the duration is the result of the hiding, so it is important for the semantic structure to distinguish the hiding from its result.

WG also offers solutions to a range of other problems of semantics; for example, it includes the non-standard version of quantification sketched in section 33.4 as well as a theory of sets and a way of distinguishing distributed and joint actions (Hudson 2006: 228–32); but this discussion can merely hint at the theory's potential.

33.10 Learning and using language

Question (j) is: "How does your model relate to studies of acquisition and to learning theory?" A central tenet of WG is that the higher levels of language are learned rather than innate, and that they are learned with the help of the same mechanisms as are available for other kinds of knowledge-based behavior. (In contrast, WG makes no claims about how the acoustics and physiology of speech develop.) This tenet follows from the claim that language is part of the general cognitive network, but it is supported by a specific proposal for how such learning takes place (Hudson 2006: 52–9), which in turn is based on a general theory of processing. The theories of learning and processing build on the basic idea of WG that language is a network, so they also provide further support for this idea.

The main elements in the WG theory of processing are activation and node-creation. As in all network models of cognition, the network is "active" in two senses. First, activation—which is ultimately expressed in terms of physical energy—circulates around the network as so-called "spreading activation", making some nodes and links temporarily active and leaving some of them permanently more easily re-activated than others. There is a great deal of evidence for both these effects. Temporary activation can be seen directly in brain imaging (Skipper and Small 2006), but also indirectly through the experimental technique of priming (Reisberg 2007: 257–62). Permanent effects come mainly from frequency of usage and emerge in experiments such as those which test the relative "availability" of words (Harley 1995: 146–8). The two kinds of change are related because temporary activation affects nodes differently according to their permanent activation level. Moreover, because there is no boundary around language, activation spreads freely between language and non-language, so the "pragmatic context" influences the way in which we interpret utterances (e.g., by guiding us to intended referents).

The second kind of activity in the network consists of constant changes in the fine details of the network's structure through the addition (and subsequent loss) of nodes and links in response to temporary activation. Many of these new nodes deal with ongoing items of experience; so (for example) as you read this page you are creating a new node for each letter-token and word-token that you read. Token nodes must be kept separate from the permanent "type nodes" in the network because the main aim of processing is precisely to match each token with some type—in other words, to classify it. The two nodes must be distinct because the match may not be perfect, so when you read *yelow*, you match it mentally with the stored word YELLOW in spite of the mis-spelling.

As for learning, WG offers two mechanisms. One is the preservation of temporary token nodes beyond their normal life expectancy of a few seconds; this might be triggered for example by the unusually high degree of activation attracted by an unfamiliar word or usage. Once preserved from oblivion, such a node would turn (logically) into a type node available for processing future token nodes. The other kind of learning is induction, which also involves the creation of new nodes. Induction is the process of spotting generalizations across nodes and creating a new super-node to express the generalization. For instance, if the network already contains several nodes which have similar links to the nodes for "wing", "beak", and "flying", a generalization emerges: wings, beaks, and flying go together; and a new node can be created which also has the same links to these three other nodes, but none of the specifics of the original nodes. Such generalizations can be expressed as a statistical correlation between the shared properties, and in a network they can be found by looking for nodes which happen to receive activation from the same range of other nodes. Induction is very different from the processing of ongoing experience, and indeed it may require down time free of urgent experience such as the break we have during sleep.

In reply to question (l) "How does your model deal with usage data?", therefore, the WG theory of learning fits comfortably in the "usage-based" paradigm of cognitive linguistics (Barlow and Kemmer 2000) in which language emerges in a rather messy and piecemeal way out of a child's experience, and is heavily influenced by the properties of the "usage" experienced, and especially by its frequency patterns (Bybee 2006*c*).

33.11 The social context

Question (i) is: "Does your model take sociolinguistic phenomena into account?" The answer to this question is probably more positive for WG than for any other theory of language structure. As explained in section 33.1, sociolinguistics has long been one of my interests—indeed, this interest predates the start of WG—and I have always tried to build some of the more relevant findings of sociolinguistics into my ideas about language structure and cognition.

One of the most relevant conclusions of sociolinguistics is that the social structures to which language relates are extremely complex, and may not be very different in complexity from language itself. This strengthens the case, of course, for the WG claim that language uses the same cognitive resources as we use for other areas of life, including our social world—what we might call "I-society", to match "I-language". The complexity of I-society lies partly in our classification of people and their permanent relations (through kinship, friendship, work, and so on); and partly in our analysis of social interactions, where we negotiate subtle variations on the basic relations of power and solidarity. It is easy to find parallels with language; for example, our permanent classification of people is similar to the permanent classification of word-types, and the temporary classification of interactions is like our processing of word-tokens.

Another link to sociolinguistics lies in the structure of language itself. Given the three-level architecture (section 33.5), language consists of sounds, forms, and words, each of which has various properties including some "social" properties. Ignoring sounds, forms are seen as a kind of action and therefore inherit (inter alia) a time and an actor—two characteristics of social interaction. Words, on the other hand, are symbols, so they too inherit interactional properties including an addressee, a purpose, and (of course) a meaning (Hudson 2006: 218). These inherited properties provide important "hooks" for attaching sociolinguistic properties which otherwise have no place at all in a model of language. To take a very elementary example, the form {bonny} has the property of being typically used by a Scot—a fact which must be part of I-language if this includes an individual's knowledge of language. Including this kind of information in a purely linguistic model is a problem for which most theories of language structure offer no solution at all, and cannot offer any solution because they assume that I-language is separate from other kinds of knowledge. In contrast, WG offers at least the foundations of a general solution as well as some reasonably well-developed analyses of particular cases (Hudson 1997*a*; 2007*a*; 2006: 246–8). To return to the example of {bonny}, the WG analysis in Figure 33.13 shows that its inherited "actor" (i.e., its speaker) isa Scot—an element in social structure (I-society), and not a mere uninterpreted diacritic.

33.12 Similarities and differences across space and time

Since WG is primarily a theory of I-language (section 33.2) it might not seem relevant to question (g): "How does your model account for typological diversity and

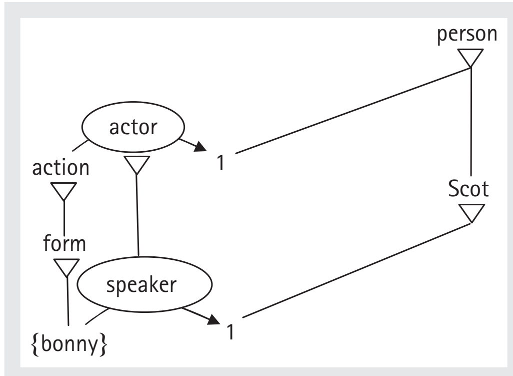


Figure 33.13. The form {bonny} is typically used by a Scot

universal features of human languages?" or (h): "How is the distinction synchrony vs. diachrony dealt with?". Typology and historical linguistics have traditionally been approached as studies of the E-language of texts and shared language systems. Nevertheless, it is individuals who change languages while learning, transmitting, and using them, so I-language holds the ultimate explanation for all variation within and between languages.

The answers to questions (g) and (h) rest on the answer to question (k): "How does your model generally relate to variation?" Variation is inherent in the WG model of I-language, partly because each individual has a different I-language but more importantly because each I-language allows alternatives to be linked to different social contexts (section 33.11). Such variation applies not only to lexical items like BONNY in relation to its synonyms, but also to phonological, morphological, and syntactic patterns—the full range of items that have been found to exhibit "inherent variability" (e.g., Labov 1969; Hudson 1996: 144–202). Moreover, variation may involve categories which range from the very specific (e.g., BONNY) to much more general patterns of inflectional morphology (e.g., uninflected 3rd-singular present verbs in English) or syntax (e.g., multiple negation). These more general patterns of social variation emerge in the network as correlations between social and linguistic properties, so learners can induce them by the same mechanisms as the rest of the grammar (section 33.10).

Returning to the two earlier questions, then, the distinction between synchrony and diachrony is made within a single I-language whenever the social variable of age is invoked, because language change by definition involves variation between the language of older and younger people and may be included in the I-language of either or both generations. However, this analysis will only reveal the ordinary speaker's understanding of language change, which may not be accurate; for example, younger speakers may induce slightly different generalizations from older speakers without being at all aware of the difference. One of the major research questions in this area is whether this "restructuring" is gradual or abrupt, but usage-based learning (section 33.10) strongly predicts gradual change because each generation's I-language is based closely on that of the previous generation. This does indeed appear to be the case with one of the test cases for the question, the development of the modern English auxiliary system (Hudson 1997*b*). As for the other question, diversity among languages must derive from the theory of change because anything which can change is a potential source of diversity. Conversely, anything which cannot change because it is essential for language must also be universal. These answers follow from the WG mechanisms for inducing generalizations.

Equally importantly, though, the same mechanisms used in such variation of individual features allow us to induce the large-scale categories that we call "languages" or "dialects", which are ultimately based, just like all other linguistic categories, on correlations among linguistic items (e.g., *the* correlates with *cup* in contrast with *la* and *tasse*) and between these and social categories. These correlations give rise to general categories such as "English word" (or "English linguistic unit", as in Figure 33.5) which allow generalizations about the language. These language-particular categories interact, thanks to multiple inheritance, with language-neutral categories such as word classes, so a typical English word such as *cup* inherits some of its properties from "English word" and others from "noun"—see Figure 33.14. The result is a model of bilingualism (Hudson 2006: 239–46) which accommodates any degree of separation or integration of the languages and any degree of proficiency, and which explains why code-mixing within a sentence is both possible and also constrained by the grammars of both languages (Wei 2006). The same model also offers a basis for a theory about how one language can influence another within a single I-language (and indirectly, in the entire E-language).

The one area of typological research where WG has already made a contribution is word order. Typological research has found a strong tendency for languages to minimize "dependency distance"—the distance between a word and the word on which it depends (e.g., Hawkins 2001), a tendency confirmed by research in psycholinguistics (Gibson 2002), and corpus linguistics (Ferrer i Cancho 2004, Collins 1996). The notion of "dependency distance" is easy to capture in a dependency-based syntactic theory such as WG, and the theory's psychological orientation suggests a research program in psycholinguistic typology. For example, it is easy to explain the popularity of SVO and similar "mixed" orders in other phrase types as a way of reducing the number of dependents that are separated from the phrase's head; thus in SVO order, both S and O are adjacent to V, whereas in both VSO and SOV one of these dependents is separated from V (Hudson 2006: s161). However, this explanation also implies that languages with different word orders may tend to make different demands on their users, when measured in terms of average dependency distances in comparable styles. Results so far suggest that this is in fact

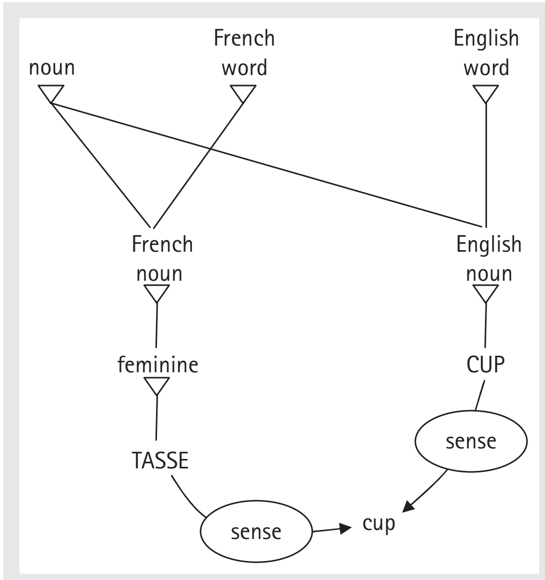


Figure 33.14. French TASSE and English CUP share a word class and a meaning

the case—for instance, average distances in Mandarin are much greater than those in English, and other languages have intermediate values (Liu et al. 2008).

What, then, does WG offer a working descriptive linguist? What it does not offer is a check-list of universal categories to be "found" in every language. The extent to which different languages require the same categories is an empirical research question, not a matter of basic theory. What it does offer is a way of understanding the structure of language in terms of general psychological principles. However, it is also important to stress that the theory has evolved over several decades of descriptive work, mostly but not exclusively on English, and dealing with a wide range of topics—in morphology, syntax, and semantics; concerning language structure, psycholinguistics, and sociolinguistics; and in bilingual as well as monolingual speech. I believe the theoretical basis provides a coherence, breadth, and flexibility which are essential in descriptive work.

This page intentionally left blank

References

Aarts, Bas (2007) *Syntactic Gradience: The Nature of Grammatical Indeterminacy*. Oxford: Oxford University Press.

Abdulaev, Yalichin G., and Michael I. Posner (1997) Time-course of activating brain areas in generating verbal associations. *Psychological Science* 8: 56–9.

Abeillé, Anne (ed.) (2003) *Treebanks*. Dordrecht: Kluwer Academic.

Abels, Klaus (2007) Towards a restrictive theory of (remnant) movement. In Jeroen Van Craenenbroeck (ed.), *Linguistic Variation Yearbook 2007*, pp. 53–120. Amsterdam: Benjamins.

Abney, Steven P. (1987) The English noun phrase in its sentential aspect. PhD dissertation, MIT.

Aboh, Enoch (2004) *The Morphosyntax of Complement-Head Sequences*. New York: Oxford University Press.

Aboh, Enoch (2007) Information structuring begins with numeration. Amsterdam Center for Language and Communication (ACLC), University of Amsterdam.

Adamson, H. D. (1992) Social and processing constraints on relative clauses. *American Speech* 67: 123–33.

Adger, David (2003) *Core Syntax*. Oxford: Oxford University Press.

Ágel, Vilmos (2000) *Valenztheorie*. Tübingen: Narr.

Ágel, Vilmos (2001) Gegenwartsgrammatik und Sprachgeschichte: methodologische Überlegungen am Beispiel der Serialisierung im Verbalkomplex. *Zeitschrift für germanistische Linguistik* 29: 319–31.

Ágel, Vilmos (2004) Phraseologismus als (valenz)syntaktischer Normalfall. In Kathrin Steyer (ed.), *Wortverbindungen—mehr oder weniger fest. Jahrbuch 2003 des Instituts für Deutsche Sprache*, pp. 65–85. Berlin: de Gruyter.

Ágel, Vilmos (2006) (Nicht)Flexion des Substantiv(s) Neue Überlegungen zum finiten Substantiv. *Zeitschrift für germanistische Linguistik* 34: 286–327.

Ágel, Vilmos, Ludwig M. Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans-Jürgen Heringer, and Henning Lobin (eds.) (2003) *Handbook of Dependency and Valency*, vol. 1. Berlin: de Gruyter.

Ágel, Vilmos, Ludwig M. Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans-Jürgen Heringer, and Henning Lobin (eds.) (2006) *Handbook of Dependency and Valency*, vol. 2. Berlin: de Gruyter.

Ahern, Aoife, and Manuel Leonetti (2004) The Spanish subjunctive: procedural semantics and pragmatic inference. In R. Márquez-Reiter and M.E. Placencia (eds.), *Current Trends in the Pragmatics of Spanish*, pp. 35–56. Amsterdam: Benjamins.

Aijmer, Karin (2002) *English Discourse Particles: Evidence from a Corpus*. Amsterdam: Benjamins.

Aikhenvald, Alexandra Y. (2003) *A Grammar of Tariana, from Northwest Amazonia*. Cambridge: Cambridge University Press.

Aissen, Judith (1999) Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17: 673–711.

Aissen, Judith (2003) Differential object marking: iconicity vs. economy. *Natural Language and Linguistic Theory* 21: 435–83.

Ajdukiewicz, Kazimierz (1935) Die syntaktische Konnexität. *Studia Philosphica* 1: 1–27. Translated in Storrs McCall (ed.), *Polish Logic: 1920–1939*, pp. 207–231. Oxford: Oxford University Press, 1967.

Akhtar, Nameera, and Michael Tomasello (1997) Young children's productivity with word order and verb morphology. *Developmental Psychology* 33: 952–65.

Aksu-Koç, Ayhan, and Didem Mersin Alici (2000) Understanding sources of beliefs and marking of uncertainty: the child's theory of evidentiality. In Eve V. Clark (ed.), *Proceedings of the 30th Annual Child Language Research Forum*, pp. 123–30. Stanford, CA: Stanford University Press.

Alishahi, Afra, and Suzanne Stevenson (2008) A computational model for early argument structure acquisition. *Cognitive Science* 32: 5.

Allen, Cynthia (1995) *Case-Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Clarendon Press.

Alon, Uri (2003) Biological networks: the tinkerer as an engineer.*Science* 301: 1866–7.

Alon, Uri (2007) Simplicity in biology. *Nature* 446: 497.

Alsina, Alex (1996) *The Role of Argument Structure in Grammar*. Stanford, CA: CSLI Publications.

Altenberg, Bengt (1998) On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. Cowie (ed.), *Phraseology: Theory, Analysis and Applications*, pp. 101–22. Oxford: Oxford University Press.

Altenberg, Bengt, and Mats Eeg-Olofsson (1990) Phraseology in spoken English. In Jan Aarts and Willem Meijs (eds.), *Theory and Practice in Corpus Linguistics*, pp. 1–26. Amsterdam: Rodopi.

Ambridge, Ben, Caroline Rowland, and Julian Pine (2008) Is structure dependence an innate constraint? New experimental evidence from chilren's complex-question production. *Cognitive Science* 32: 222–55.

Amha, Azeb (2001) *The Maale Language*. Leiden: University of Leiden, Research School of Asian, African, and Amerindian Studies.

Amiot, Dany (2005) Between compounding and derivation: elements of word formation corresponding to prepositions. In Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer (eds.), *Morphology and its Demarcations*, pp. 81–96. Amsterdam: Benjamins.

Andersen, Roger W. (1979) Explaining Schumann's Pidginization Hypothesis. *Language Learning* 29, 1: 105–19.

Amundson, Ron (2005) *The Changing Role of the Embryo in Evolutionary Thought*. Cambridge: Cambridge University Press.

Anderson, John M. (1971) *The Grammar of Case: Towards a Localistic Theory*. Cambridge: Cambridge University Press.

Anderson, John R. (1978) Arguments concerning representations for mental imagery. *Psychological Review* 85: 249–77.

Anderson, John R. (1993) *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

Anderson, Stephen R. (1992) *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. (1999) A formalist's reading of some functionalist work in syntax. In Darnell et al. (1999: 111–35).

Andersson, Peter (2006) Swedish *må* and the (de)grammaticalization debate. TS, Institutionen för svenska språket, Göteborgs Universitet.

Andrews, Avery (1990) Unification and morphological blocking. *Natural Language and Linguistic Theory* 8: 507–57.

Andrews, Avery (2007) The major functions of the noun phrase. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 1: *Clause Structure*, 2nd edn, pp. 132–223. Cambridge: Cambridge University Press.

Anttila, Arto (2002) Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20: 1–42.

Archangeli, Diana, and D. Terence Langendoen (1997) *Optimality Theory: An Overview*. Oxford: Blackwell.

Ariel, Mira (2004) Most. *Language* 80: 658–706.

Ariel, Mira (2006) A 'just that' lexical meaning for *most*. In von Heusinger and Turner (2006: 49–91).

Arka, I. Wayan, and J. Kosmas (2005) Passive without passive morphology? Evidence from Manggarai. In I. W. Arka and M. D. Ross (eds.), *The Many Faces of Austronesian Voice Systems: Some New Empirical Studies*, pp. 87–117. Canberra: Pacific Linguistics.

Armstrong, David F., William C. Stokoe, and Sherman E. Wilcox (1995) *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

Armstrong, David F., and Sherman E. Wilcox (2007) *The Gestural Origin of Language*. Oxford: Oxford University Press.

Armstrong, Elizabeth (2001) Connecting lexical patterns of verb usage with discourse meanings in aphasia. *Aphasiology* 15: 1029–46.

Arnold, Doug, and Evita Linardaki (2007) Linguistic constraints in LFG-DOP. In Butt and King (2007: 66–86).

Arnon, Inbal, Neal Snider, Philip Hofmeister, T. Florian Jaeger, and Ivan A. Sag (forthcoming) Cross-linguistic variation in a processing account: the case of multiple wh-questions. *Berkeley Linguistics Society* 32.

Aronoff, Mark (1976) *Word Formation in Gernative Grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark (1994) *Morphology By Itself: Stems And Inflectional Classes*. Cambridge, MA: MIT Press.

Aronoff, Mark, and Janie Rees-Miller (eds.) (2000) *The Handbook of Linguistics*. Oxford: Blackwell.

Aronoff, Mark, and Janie Rees-Miller (eds.) (2006) *The Handbook of Linguistics*, 2nd edn. Oxford: Blackwell.

Asbury, Anna, Jakub Dotlačil, Berit Gehrke, and Rick Nouwen (eds.) (2008) *Syntax and Semantics of Spatial P*. Amsterdam: Benjamins.

Ashton, Ethel O. (1944) *Swahili Grammar (including Intonation)* London: Longman.

Asudeh, Ash (2004) Resumption as resource management. PhD thesis, Stanford University.

Asudeh, Ash (2006) Direct compositionality and the architecture of LFG. In Butt and King (2006: 363–87).

Asudeh, Ash, and Richard Crouch (2002) Glue semantics for HPSG. In Frank van Eynde, Lars Hellan, and Dorothee Beermann (eds.), *Proceedings of the 8th International HPSG Conference*, pp. 1–19. Stanford, CA: CSLI Publications.

Asudeh, Ash, Mary Dalrymple, and Ida Toivonen (2008) Constructions with lexical integrity: templates as the lexicon-syntax interface. In Butt and King (2008: 68–88).

Atkinson, Dwight (1992) The evolution of medical research writing from 1735 to 1985: the case of the *Edinburgh Medical Journal*. *Applied Linguistics* 13: 337–74.

Atkinson, Dwight (1996) The philosophical transactions of the Royal Society of London, 1675–1975: a sociohistorical discourse analysis. *Language in Society* 25: 333–71.

Atkinson, Dwight (1999) *Scientific Discourse in Sociohistorical Context: The Philosophical Transactions of the Royal Society of London, 1675–1975*. Hillsdale, NJ: Erlbaum.

Atkinson, Richard C., and Richard M. Shiffrin (1968) Human memory: a proposed system and its control processes. In Kenneth W. Spence and Janet T. Spence (eds.), *The Psychology of Learning and Motivation*, vol. 2, pp. 89–195. New York: Academic Press.

Atlas, Jay David (1977) Negation, ambiguity, and presupposition. *Linguistics and Philosophy* 1: 321–36.

Atlas, Jay David (1979) How linguistics matters to philosophy: presupposition, truth, and meaning. In C.-K. Oh and D. A. Dinneen (eds.), *Syntax and Semantics*, vol. 11, pp. 265–81. New York: Academic Press.

Atlas, Jay David (2005) *Logic, Meaning, and Conversation*. Oxford: Oxford University Press.

Atlas, Jay David (2006) Remarks on F. Recanati's *Literal Meaning*. MS.

Atlas, Jay David, and Stephen C. Levinson (1981) It-clefts, informativeness and logical form: radical pragmatics. In P. Cole (ed.), *Radical Pragmatics*, pp. 1–61. London: Academic Press.

Au, Terry Kit-Fong (1983) Chinese and English counterfactuals: the Sapir-Whorf hypothesis revisited. *Cognition* 15: 155–87.

Au, Terry Kit-Fong (1984) Counterfactuals: in reply to Alfred Bloom. *Cognition* 17: 289–302.

Austin, Gavin (2008) The indefinite article in Sinhala: implications for a constraint-based analysis. *Proceedings of the 6th International Conference of Cognitive Science* (ICCS 2008, Yonsei University, Korea).

Austin, John (1962) *How to Do Things with Words*. Cambridge, MA: Harvard University Press.

Baayen, R. Harald (1992) Quantitative aspects of morphological productivity. In Geert Booij and Jaap van Marle (eds.), *Yearboook of Morphology 1991*, pp. 109–50. New York: Springer.

Baayen, R. Harald (2008a) Corpus linguistics in morphology: morphological productivity. In Anke Lüdeling, M. Kyto, and T. McEnery (eds.), *Handbook of Corpus Linguistic*. Berlin: de Gruyter.

Baayen, R. Harald (2008b) *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.

Baayen, R. Harald, Ton Dijkstra, and Rob Schreuder (1997) Singulars and plurals in Dutch: evidence for a parallel dual route model. *Journal of Memory and Language* 36: 94–117.

Baayen, R. Harald, and Robert Schreuder (eds.) (2003) *Morphological Structure in Language Processing*. Berlin: Mouton de Gruyter.

BACH, EMMON (2004) Linguistic universals and particulars. In Piet van Sterkenburg (ed.), *Linguistics Today: Facing a Greater Challenge*, pp. 47–60. Amsterdam: Benjamins.

BACH, KENT (1994) Semantic slack: what is said and more. In S. L. Tsohatzidis (ed.), *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, pp. 267–91. London: Routledge.

BACH, KENT (2004a) Minding the gap. In C. Bianchi (ed.), *The Semantics/Pragmatics Distinction*, pp. 27–43. Stanford, CA: CSLI Publications.

BACH, KENT (2004b) Pragmatics and the philosophy of language. In Horn and Ward (2006: 463–87).

BACH, KENT (2006a) The excluded middle: semantic minimalism without minimal propositions. *Phillosophy and Phenomenological Research* 73: 435–42.

BACH, KENT (2006b) The top ten misconceptions about implicature. In B. J. Birner and G. Ward (eds.), *Drawing the Boundaries of Meaning*. Amsterdam: Benjamins.

BAERMAN, MATTHEW, and DUNSTAN BROWN (2005) Case syncretism. In *WALS*, pp. 118–21.

BAERTSCH, KAREN (2002. An Optimality-theoretic approach to syllable structure: the Split Margin Hierarchy. Doctoral dissertation, Indiana University, Bloomington.

BAILEY, DAVID, NANCY CHANG, JEROME FELDMAN, and SRINI NARAYANAN (1998) Extending embodied lexical development. In Morton Ann Gernsbacher and Sharon J. Derry (eds.), *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society* (COGSCI-98), pp. 84–90. Hillsdale, NJ: Erlbaum.

BAKER, CHARLOTTE, and DENNIS COKELY (1980) *American Sign Language: A Teacher's Resource Text on Grammar and Culture*. Silver Spring, MD: T.J. Publishers.

BAKER, MARK C. (1985) The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–416.

BAKER, MARK C. (1988) *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

BAKER, MARK C. (1996) *The Polysynthesis Parameter*. Oxford: Oxford University Press.

BAKER, MARK C. (2001) *The Atoms of language: The Mind's Hidden Rules of Grammar*. New York: Basic Books.

BAKER, MARK C. (2003) *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.

BAKER, MARK C. (2008a) *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

BAKER, MARK C. (2008b) The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, pp. 351–74. Amsterdam: Benjamins.

BAKER, MARK C., KYLE JOHNSON, and IAN ROBERTS (1989) Passive arguments raised. *Linguistic Inquiry* 20: 219–51.

BAKER, MARK C., and JAMES MCCLOSKEY (2007) On the relationship of typology to theoretical syntax. *Linguistic Typology* 11: 285–96.

BAKER, PHILIP, and ANAND SYEA (eds.) (1996) *Changing Meanings, Changing Functions: Papers Relating to Grammaticalization in Contact Languages*. London: University of Westminster Press.

BAKOVIĆ, ERIC (1998) Unbounded stress and factorial typology. In Ron Artstein and Madeleine Holler (eds.), *Ruling Papers 1: Working Papers from Rutgers University*, pp. 15–28. New Brunswick, NJ: Dept of Linguistics, Rutgers University.

BAKOVIĆ, ERIC (2005) Antigemination, assimilation and the determination of identity. *Phonology* 22: 279–315.

Bar-Hillel, Yehoshua (1953) A quasi-arithmetical notation for syntactic description. *Language* 29: 47–58.
Barbieri, Maria Silvia, and Elisabetta Bascelli (2000) Comprehension of epistemic and deontic verbal modalities during development. *Giornale italiano di psicologia* 27: 529–48.
Barbosa, Pilar, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (1998) *Is the Best Good Enough? Optimality and Competition in Syntax*. Cambridge, MA: MIT Press.
Barlow, Michael, and Suzanne Kemmer (eds.) (2000) *Usage-Based Models of Language*. Stanford, CA: CSLI Publications.
Baron-Cohen, Simon (1995) *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
Baron-Cohen, Simon (2000) The evolution of a theory of mind. In Michael Corballis and Stephen E. G. Lea (eds.), *The Descent of Mind: Psychological Perspectives on Hominid Evolution*, pp. 261–77. Oxford: Oxford University Press.
Barsalou, Lawrence W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–660.
Barss, Andrew, and Howard Lasnik (1986) A note on anaphora and double objects. *Linguistic Inquiry* 17: 347–54.
Bartlett, Frederick (1932) *Remembering*. Cambridge: Cambridge University Press.
Bates, Elizabeth, Laura Benigni, Inge Bretherton, Luigi Camioni, and Virginia Volterra (1979) *The Emergence of Symbols: Cognition and Communication in Infancy*. New York: Academic Press.
Bates, Elizabeth, Luigi Camioni, and Virginia Volterra (1975) The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly* 21: 205–26.
Bates, Elizabeth, and Brian MacWhinney (1988) What is functionalism? *Papers and Reports on Child Language Development* 27: 137–52.
Battig, W. F., and W. E. Montague. (1969) Category norms for verbal items in 56 categories: a replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph* 80.3.2: 1–46.
Battison, Robbin (1978) *Lexical Borrowing in American Sign Language*. Silver Spring, MD: Linkstok Press.
Bauer, Laurie (2001) *Morphological Productivity*. Cambridge: Cambridge University Press.
Beard, Robert (1994) Lexeme-morpheme base morphology. In Ronald Asher (ed.), *Encyclopedia of Language and Linguistics*, pp. 2137–40. Oxford: Pergamon.
Beckner, Clay, and Joan Bybee (2009) A usage-based account of constituency and reanalysis. *Language Learning* 59, Supplement 1.
Bell, Alan (1978) Syllabic consonants. In J. H. Greenberg (ed.), *Universals of Human Language*, vol. 2: *Phonology*, pp. 153–201. Stanford, CA: Stanford University Press.
Belletti, Adriana (1990) *Generalized Verb Movement*. Turin: Rosenberg & Sellier.
Belletti, Adriana (ed.) (2004) *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3. New York: Oxford University Press.
Belletti, Adriana (2009) *Structures and Strategies*. London: Routledge.
Belloro, V. (2007) Spanish clitic doubling: a study of the syntax-pragmatics interface. PhD dissertation, University at Buffalo.
Bender, Emily, and Andreas Kathol (forthcoming) Constructional effects of *Just because . . . doesn't mean . . . Berkeley Linguistics Society* 27.

Benincà, Paola (2001) The position of topic and focus in the left periphery. In G. Cinque and G. Salvi (eds.), *Current Studies in Italian Syntax: Essays offered to Lorenzo Renzi*, pp. 39–64. Amsterdam: Elsevier North-Holland.

Benincà, Paola (2006) A detailed map of the left periphery of medieval Romance. In R. Zanuttini, H. Campos, E. Herburger, and P. Portner (eds.), *Crosslinguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture*, pp. 53–86. Washington, DC: Georgetown University Press.

Benincà, Paola, and Nicola Munaro (eds.) (forthcoming) *Mapping the Left Periphery: The Cartography of Syntactic Structures*, vol. 5. New York: Oxford University Press.

Benincà, Paola, and Cecilia Poletto (2004) Topic, focus, and V2: defining the CP sublayers. In L. Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures*, vol. 2, pp. 52–75. New York: Oxford University Press.

Benua, Laura (1997) Transderivational identity: phonological relations between words. Doctoral dissertation, University of Massachusetts, Amherst. Available on the Rutgers Optimality Archive, ROA 259, http://roa.rutgers.edu

Benz, Anton, Gerhard Jäger, and Robert van Rooy (2006) *Game Theory and Pragmatics*. Basingstoke: Palgrave Macmillan.

Berbeira Gardón, José Luis (1996) *Los verbos modales ingleses: estudio semántico-pragmático*. Cadiz: Servicio de Publicaciones.

Bergen, Benjamin (2004) The psychological reality of phonaesthemes. *Language* 80: 290–311.

Bergen, Benjamin, and Nancy Chang (2005) Embodied construction grammar in simulation-based language understanding. In Jan-Olastman and Miriam Fried (eds.), *Construction Grammars: Cognitive Grounding and Theoretical Extensions*, pp. 147–90. Amsterdam: Benjamins.

Berman, Judith (2003) *Clausal Syntax of German*. Stanford, CA: CSLI Publications.

Berman, R. A., and D. I. Slobin (1994) *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, NJ: Erlbaum.

Besnier, Niko (1988) The linguistic relationships of spoken and written Nukulaelae registers. *Language* 64: 707–36.

Bianchi, Valentina (2006) On the syntax of personal arguments. *Lingua* 116: 2023–67.

Bianco, Maria T. (1996) *Valenzlexikon deutsch-italienisch* (2 vols). Heidelberg: Groos.

Biber, Douglas (1986) Spoken and written textual dimensions in English: resolving the contradictory findings. *Language* 62: 384–414.

Biber, Douglas (1988) *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas (1989) The typology of English texts. *Linguistics* 27: 3–43.

Biber, Douglas (1992) On the complexity of discourse complexity: a multidimensional analysis. *Discourse Processes* 15: 133–63. Reprinted in Conrad and Biber (2001: 215–40).

Biber, Douglas (1993) Representativeness in corpus design. *Literary and Linguistic Computing* 8: 1–15.

Biber, Douglas (1995) *Dimensions of Register Variation: A Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Biber, Douglas (2001) Dimensions of variation among eighteenth-century speech-based and written registers. In Conrad and Biber (2001: 200–214).

Biber, Douglas (2006) *University Language: A Corpus-Based Study of Spoken and Written Registers*. Amsterdam: Benjamins.

Biber, Douglas (2008) Corpus-based analyses of discourse: dimensions of variation in conversation. In V. Bhatia, J. Flowerdew, and R. Jones (eds.), *Advances in Discourse Studies*, pp. 100–114. London: Routledge.

Biber, Douglas, and Federica Barbieri (2007) Lexical bundles in university spoken and written registers. *English for Specific Purposes* 26: 263–86.

Biber, Douglas, and Jená Burges (2000) Historical change in the language use of women and men: gender differences in dramatic dialogue. *Journal of English Linguistics* 28: 21–37.

Biber, Douglas, Ulla Connor, and Thomas A. Upton (eds.) (2007) *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam: Benjamins.

Biber, Douglas, and Susan Conrad (1999) Lexical bundles in conversation and academic prose. In H. Hasselgard and S. Oksefjell (eds.), *Out of Corpora: Studies in honor of Stig Johansson*, pp. 181–9. Amsterdam: Rodopi.

Biber, Douglas, and Susan Conrad (forthcoming) *Register and Genre Variation*. Cambridge: Cambridge University Press.

Biber, Douglas, Susan Conrad, and Viviana Cortes (2004) *If you look at . . .*: lexical bundles in university teaching and textbooks. *Applied Linguistics* 25: 371–405.

Biber, Douglas, Susan Conrad, and Randi Reppen (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, Douglas, Susan Conrad, Randi Reppen, Pat Byrd, and Marie Helt (2002) Speaking and writing in the university: a multi-dimensional comparison. *TESOL Quarterly* 36: 9–48.

Biber, Douglas, Mark Davies, James K. Jones, and Nicole Tracy-Ventura (2006) Spoken and written register variations in Spanish: a multi-dimensional analysis. *Corpora* 1: 7–38.

Biber, Douglas, and Edward Finegan (1989) Drift and the evolution of English style: a history of three genres. *Language* 65: 487–517.

Biber, Douglas, and Edward Finegan (1994a) Multi-dimensional analyses of authors' styles: some case studies from the eighteenth century. In D. Ross and D. Brink (eds.), *Research in Humanities Computing 3*, pp. 3–17. Oxford: Oxford University Press.

Biber, Douglas, and Edward Finegan (1994b) Intra-textual variation within medical research articles. In N. Oostdijk and P. de Haan (eds.), *Corpus-Based Research into Language*, pp. 201–22. Amsterdam: Rodopi. Reprinted in Conrad and Biber (2001: 108–23).

Biber, Douglas, and Edward Finegan (1997) Diachronic relations among speech-based and written registers in English. In T. Nevalainen and L. Kahlas-Tarkka (eds.), *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, pp. 253–75. Helsinki: Société Neophilologique. Reprinted in Conrad and Biber (2001: 66–83).

Biber, Douglas, and Mohamed Hared (1992) Dimensions of register variation in Somali. *Language Variation and Change* 4: 41–75.

Biber, Douglas, and Mohamed Hared (1994) Linguistic correlates of the transition to literacy in Somali: language adaptation in six press registers. In D. Biber and E. Finegan (eds.), *Sociolinguistic Perspectives on Register*, pp. 182–216. New York: Oxford University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan (1999) *Longman Grammar of Spoken and Written English*. London: Longman.

Biber, Douglas, and James K. Jones (2005) Merging corpus linguistic and discourse analytic research goals: discourse units in biology research articles. *Corpus Linguistics and Linguistic Theory* 1: 151–82.

Biber, Douglas, YouJin Kim, and Nicole Tracy-Ventura (forthcoming) A corpus-driven approach to comparative phraseology: lexical bundles in English, Spanish, and Korean. In S. Iwasaki (ed.), *Japanese/Korean Linguistics*, vol. 17. Stanford, CA: CSLI Publications.

Biber, Douglas, and Jerry Kurjian (2007) Towards a taxonomy of web registers and text types: a multi-dimensional analysis. In M. Hundt, N. Nesselhauf, and C. Biewer (eds.), *Corpus Linguistics and the Web*, pp. 109–32. Amsterdam: Rodopi.

Biber, Douglas, and Nicole Tracy-Ventura (2007) Dimensions of register variation in Spanish. In G. Parodi (ed.) *Working with Spanish Corpora*, pp. 54–89. London: Continuum.

Bickel, Balthasar (2000) Grammar and social practice: on the role of 'culture' in linguistic relativity. In S. Niemeier and R. Dirven (eds.), *Evidence for Linguistic Relativity*, pp. 161–91. Amsterdam: Benjamins.

Bickel, Balthasar (2007) Typology in the 21st century: major current developments. *Linguistic Typology* 11: 239–51.

Bickerton, Derek (1981) *Roots of Language*. Ann Arbor, MI: Karoma.

Bickerton, Derek (1990) *Language and Species*. Chicago: University of Chicago Press.

Bickerton, Derek (forthcoming) Recursion: core of complexity or artifact of analysis? In M. Shibatani and T. Givón (eds.), *The Genesis of Syntactic Complexity*. Amsterdam: Benjamins.

Bickerton, Derek, and Carol Odo (1976) Change and variation in Hawaiian English, vol. 1: The pidgin. MS, University of Hawaii.

Birner, Betty J., and Gregory Ward (1998) *Information Status and Noncanonical Word Order in English*. Amsterdam: Benjamins.

Bisang, Walter (1996) Areal typology and grammaticalization: processes of grammaticalization based on nouns and verbs in East and Mainland South East Asian languages. *Studies in Language* 20: 519–97.

Bisang, Walter, and Peter Rinderknecht (eds.) (1991) *Von Europa bis Ozeanien: von der Antonymie zum Relativsatz*. Zürich: Universität Zürich.

Bishop, Judith, and Janet Fletcher (2005) Intonation in six dialects of Bininj Gunwok. In Sun-Ah Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, pp. 331–61. Oxford: Oxford University Press.

Bittner, Maria, and Kenneth Hale (1996) Ergativity: toward a theory of a heterogeneous class. *Linguistic Inquiry* 27: 531–604.

Black, Ezra, John Lafferty, and Salim Roukos (1992) Development and evaluation of a broad-coverage probabilistic grammar of English-language computer manuals. *Proceedings of the 30th Association of Computer Linguistics Conference* (ACL 92, Newark, Delaware), pp. 185–92.

Blackburn, Patrick, and Claire Gardent (1995) A specification language for lexical functional grammars. *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, pp. 39–44. San Francisco, CA: Morgan Kaufmann.

Blaho, Sylvia, Patrick Bye, and Martin Krämer (eds.) (2007) *Freedom of Analysis*? Berlin: Mouton de Gruyter.

Blakemore, Diane (1987) *Semantic Constraints on Relevance*. Oxford: Blackwell.

Blakemore, Diane (2002) *Relevance and Linguistic Meaning*. Cambridge: Cambridge University Press.

BLEVINS, JAMES P. (2006) Syntactic features and feature structures. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, pp. 390–93. Oxford: Elsevier.

BLEVINS, JULIETTE (2004) *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

BLOOM, ALFRED H. (1981) *The Linguistic Shaping of Thought: A Study in the Impact of Language on Thinking in China and the West*. Hillsdale, NJ: Erlbaum.

BLOOM, ALFRED H. (1984) Caution—the words you use may affect what you say: a response to Au. *Cognition* 17: 275–87.

BLOOM, LOUIS (1973) *One Word at a Time: The Use of Single Word Utterances Before Syntax*. The Hague: Mouton.

BLOOMFIELD, LEONARD (1933) *Language*. New York: Holt, Rinehart & Winston.

BLUMSTEIN, SHEILA E., and WILLIAM MILBERG (1983) Automatic and controlled processing in speech/language deficits in aphasia. In *Symposium on Automatic Speech*. Minneapolis, MN: Academy of Aphasia.

BLUTNER, REINHARD (2000) Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17: 189–216.

BLUTNER, REINHARD (2004) Pragmatics and the lexicon. In Horn and Ward (2004: 488–514).

BLUTNER, REINHARD, and HENK ZEEVAT (eds.) (2003) *Optimality Theory and Pragmatics*. Basingstoke: Palgrave Macmillan.

BOAS, HANS C. (2003) *A Constructional Approach to Resultatives*. Stanford, CA: CSLI Publications.

BOAS, HANS C. (2005) Semantic frames as interlingual representations for multilingual lexical databases. *International Journal of Lexicography* 18: 445–78.

BOBALJIK, JONATHAN D. (1999) Adverbs: the hierarchy paradox. *Glot International* 4: 27–8.

BOBALJIK, JONATHAN D., and PHIL BRANIGAN (2007) Eccentric agreement and multiple case checking. In Alana Johns, Diane Massam, and Juvenal Ndayiragije (eds.), *Ergativity*. Dordrecht: Springer.

BOBALJIK, JONATHAN D., and HÖSKULDUR THRÁINSSON (1998) Two heads aren't always better than one. *Syntax* 1: 37–71.

BOCCI, GIULIANO (2004) Contrastive focalisation on topics and preverbal subjects in Italian. *Rivista di grammatica generativa* 29: 3–60.

BOCCI, GIULIANO (2009) On syntax and prosody in Italian. PhD dissertation, University of Siena.

BOCHNER, HARRY (1984) Inflection within derivation. *Linguistic Review* 3: 411–21.

BOD, RENS (1992) A computational model of language performance: data-oriented parsing. *Proceedings COLING 1992* (Nantes), pp. 855–9.

BOD, RENS (1998) *Beyond Grammar: An Experienced-Based Theory of Language*. Stanford, CA: CSLI Publications.

BOD, RENS (1999) Context-sensitive spoken dialogue processing with the DOP Model. *Natural Language Engineering* 5: 309–23.

BOD, RENS (2001) Sentence memory: storage vs. computation of frequent sentences. *CUNY Conference on Sentence Processing 2001* (Philadelphia).

BOD, RENS (2006a) Exemplar-based syntax: how to get productivity from examples. *Linguistic Review* 23: 291–320.

BOD, RENS (2006b) An all-subtrees approach to unsupervised parsing. *Proceedings ACL-COLING 2006* (Sydney), pp. 865–72.

BOD, RENS (2007a) Is the end of supervised parsing in sight? *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition* (ACL 2007, Prague), pp. 400–407.

BOD, RENS (2007b) A linguistic investigation into U-DOP.*Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition* (ACL 2007), pp. 1–8.

BOD, RENS (2008) Constructions at work or at rest? *Cognitive Linguistics* 20: 129–34.

BOD, RENS (forthcoming) From exemplar to grammar: a probabilistic analogy-based model of language learning. *Cognitive Science*.

BOD, RENS, JENNIFER HAY, and STEFANIE JANNEDY (eds.) (2003) *Probabilistic Linguistics*. Cambridge, MA: MIT Press.

BOD, RENS, and RONALD KAPLAN (1998) A probabilistic corpus-driven model for lexical-functional analysis. *Proceedings ACL-COLING 1998* (Montreal).

BOD, RENS, and RONALD KAPLAN (2003) A DOP model for lexical-functional representations. In Bod et al. (2003), pp. 211–232.

BOD, RENS, REMKO SCHA, and KHALIL SIMA'AN (eds.) (2003) *Data-Oriented Parsing*. Stanford, CA: CSLI Publications.

BOECKX, CEDRIC (2006) *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.

BOECKX, CEDRIC (2008a) *Understanding Minimalist Syntax: Lessons from Locality in Long-Distance Dependencies*. Oxford: Blackwell.

BOECKX, CEDRIC (2008b) Elementary syntactic structures. MS, Harvard University.

BOECKX, CEDRIC (forthcoming) Approaching parameters from below. In Anna Maria di Sciullo and Cedric Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford: Oxford University Press.

BOERSMA, PAUL, and BRUCE HAYES (2001) Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32: 45–86. Available on Rutgers Optimality Archive: http://ruccs.rutgers.edu/roa.html

BOHNEMEYER, JÜRGEN, NICHOLAS J. ENFIELD, JAMES ESSEGBEY, IRAIDE IBARRETXE-ANTUÑANO, SOTARO KITA, FRIEDERIKE LÜPKE, and FELIX K. AMEKA (2007) Principles of event segmentation in language: the case of motion events. *Language* 83: 1–38.

BOLAND, ANNERIEKE (2006) *Aspect, Tense and Modality: Theory, Typology, Acquisition*. Utrecht: LOT.

BOLINGER, DWIGHT (1965) The atomization of meaning. *Language* 41: 555–73.

BOLINGER, DWIGHT (1976) Meaning and memory. *Forum Linguisticum* 1: 1–14.

BOLINGER, DWIGHT (1986) *Intonation and its Parts: Melody in Spoken English*. Stanford, CA: Stanford University Press.

BONNEMA, REMKO, RENS BOD, and REMKO SCHA (1997) A DOP model for semantic interpretation. *Proceedings ACL/EACL 1997* (Madrid), pp. 159–67. Morristown, NJ: Association for Computational Linguistics.

BOOIJ, GEERT (1993) Against split morphology. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1993*, pp. 27–49. Dordrecht: Kluwer.

BOOIJ, GEERT (1996) Inherent versus contextual inflection and the split morphology hypothesis. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1995*, pp. 1–16. Dordrecht: Kluwer.

BOOIJ, GEERT (1999a) Lexical storage and regular processes. *Brain and Behavioural Sciences* 22: 1016.

BOOIJ, GEERT (1999b) *The Phonology of Dutch*. Oxford: Oxford University Press.

Booij, Geert (2000) The phonology–morphology interface. In Lisa Cheng and Rint Sybesma (eds.), *The first Glot International State-of-the-Article Book: The Latest in Linguistics*, pp. 287–306. Berlin: Mouton de Gruyter.

Booij, Geert (2002a) Prosodic restrictions on affixation in Dutch. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2001*, pp. 183–202. Dordrecht: Kluwer.

Booij, Geert (2002b) *The Morphology of Dutch*. Oxford: Oxford University Press.

Booij, Geert (2005a) Compounding and derivation: evidence for construction morphology. In Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer (eds.), *Morphology and its Demarcations*, pp. 109–32. Amsterdam: Benjamins.

Booij, Geert (2005b) Construction-dependent morphology. *Lingue e linguaggio* 4: 31–46.

Booij, Geert (2007) *The Grammar of Words: An Introduction to Morphology*, 2nd edn. Oxford: Oxford University Press.

Booij, Geert (2009) Lexical integrity as a morphological universal: a constructionist view. In Elisabetta Magni, Sergio Scalise, and Antonietta Bisetto (eds.), *Universals of Language Today*, pp. 83–100. Dordrecht: Springer.

Booij, Geert, and Jerzy Rubach (1984) Morphological and prosodic domains in Lexical Phonology. *Phonology Yearbook* 1: 1–27.

Borensztajn, Gideon, and Willem Zuidema (2007) Bayesian model merging for unsupervised constituent labeling and grammar induction. Technical Report PP-2007–40, ILLC, University of Amsterdam.

Borensztajn, Gideon, Willem Zuidema, and Rens Bod (2008) Children's grammars grow more abstract with age: evidence from an automatic procedure for identifying the productive units of language. *Proceedings CogSci 2008* (Washington, DC).

Borer, Hagit (1984) *Parametric Syntax*. Dordrecht: Foris.

Borer, Hagit (2005) *Structuring Sense* (2 vols). Oxford: Oxford University Press.

Bornkessel, I., and M. Schlesewsky (2006a) The Extended Argument Dependency Model: a neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113: 787–821.

Bornkessel, I., and M. Schlesewsky (2006b) Generalised semantic roles and syntactic templates: a new framework for language comprehension. In Bornkessel et al. (2006: 327–53).

Bornkessel, I., M. Schlesewsky, B. Comrie, and A. Friederici (eds.) (2006) *Semantic Role Universals and Argument Linking: Theoretical, Typological and Psycholinguistic Perspectives*. Berlin: Mouton de Gruyter.

Bornkessel, I., M. Schlesewsky, and R. D. Van Valin, Jr. (2004) Syntactic templates and linking mechanisms: a new approach to grammatical function asymmetries. Poster presented at 17th CUNY Conference on Human Sentence Processing. Available on RRG website: http://linguistics.buffalo.edu/research/rrg.html

Bornkessel-Schlesewsky, I., and M. Schlesewsky (2008) Unmarked transitivity: a processing constraint on linking. In Van Valin (2008b: 413–34).

Borochovsky Bar-Aba, Esther (2003) Punctuation marks: procedural and conceptual uses. *Journal of Pragmatics* 35: 1031–48.

Boroditsky, Lera (2000) Metaphoric structuring: understanding time through spatial metaphors. *Cognition* 75: 1–28.

Boroditsky, Lera (2001) Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43: 1–22.

Boroditsky, Lera, Lauren A. Schmidt, and Webb Phillips (2003) Sex, syntax and semantics. In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind*, pp. 61–79. Cambridge, MA: MIT Press.

Borowsky, Toni (1986) Topics in the lexical phonology of English. Doctoral dissertation, University of Massachusetts, Amherst.

Bošković, Željko, and Howard Lasnik (eds.) (2007) *Minimalist Syntax: The Essential Readings*. Oxford: Blackwell.

Botha, Rudolf (2005) Pidgin languages as a putative window on language evolution. *Language and Communication* 26: 1–14.

Bouma, Gosse (2006) Unification, classical and default. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn. Oxford: Elsevier.

Bowerman, Melissa (1973) *Early Syntactic Development*. Cambridge: Cambridge University Press.

Bowerman, Melissa, and Penelope Brown (eds.) (2008) *Crosslinguistic Perspectives on Argument Structure.* Mahwah, NJ: Erlbaum.

Bowerman, Melissa, and Soonja Choi (2001) Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman and S. C. Levinson (eds.), *Language Acquisition and Conceptual Development*, pp. 475–511. Cambridge: Cambridge University Press.

Boyes-Braem, Penny (1981) Features of the handshape in American Sign Language. Thesis, University of California, Berkeley.

Brachman, Ronald J., and James G. Schmolze (1985) An overview of the KL-ONE knowledge representation system. *Cognitive Science* 9: 171–216.

Brame, Michael (1978) *Base Generated Syntax*. Seattle, WA: Noit Amrofer.

Brenier, Jason M., and Laura A. Michaelis (2005) Optimization via syntactic amalgam: syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory* 1: 45–88.

Brennan, Mary (1990) *Word Formation in British Sign Language.* Stockholm: University of Stockholm.

Brentari, Diane (1998) *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.

Bresnan, Joan (1977) Transformations and categories in syntax. In Robert E. Butts and Jaakko Hintikka (eds.), *Basic Problems in Methodology and Linguistics. Part Three of the Proceedings of the Fifth International Congress of Logic, Methodology and Philosophy of Science*, pp. 283–304. Dordrecht: Reidel.

Bresnan, Joan (1978) A realistic transformational grammar. In Morris Halle, Joan Bresnan, and George A. Miller (eds.), *Linguistic Theory and Psychological Reality*, pp. 1–59. Cambridge, MA: MIT Press.

Bresnan, Joan (ed.) (1982a) *The Mental Representation of Grammatical Relations.* Cambridge, MA: MIT Press.

Bresnan, Joan (1982b) The passive in lexical theory. In Bresnan (1982a: 3–86).

Bresnan, Joan (1982c) Control and complementation. *Linguistic Inquiry* 13: 343–434.

Bresnan, Joan (1995) Linear order, syntactic rank, and empty categories: on weak crossover. In Dalrymple et al. (1995: 241–74).

Bresnan, Joan (1997) The emergence of the unmarked pronoun: Chichewâ pronominals in Optimality Theory. In Ashlee C. Bailey, Kevin E. Moore, and Jeri L. Moxley (eds.), *Proceedings of the Berkeley Linguistics Society 23: Special Session on Syntax and Semantics in Africa*. Berkeley, CA: Berkeley Linguistics Society.

Bresnan, Joan (1998) Morphology competes with syntax: explaining typological variation in weak crossover effects. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*, pp. 59–92. Cambridge, MA: MIT Press.

Bresnan, Joan (2000) Optimal Syntax. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*, pp. 334–85. Oxford: Oxford University Press.

Bresnan, Joan (2001a) Explaining morphosyntactic competition. In Mark Baltin and Chris Collins (eds.), *Handbook of Contemporary Syntactic Theory*, pp. 11–44. Oxford: Blackwell.

Bresnan, Joan (2001b) *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler (forthcoming) *Lexical-Functional Syntax*, 2nd edn. Oxford: Wiley-Blackwell.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and R. Harald Baayen (2007) Predicting the dative alternation. In G. Bouma, I. Kraemer, and J. Zwarts (eds.), *Cognitive Foundations of Interpretation*, pp. 69–94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, Joan, Ashwini Deo, and Devyani Sharma (2007) Typology in variation: a probabilistic approach to *be* and *n't* in the *Survey of English Dialects*. *English Language and Linguistics* 11: 301–46.

Bresnan, Joan, and Jennifer Hay (2007) Gradient grammar: an effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118: 245–59.

Bresnan, Joan, and Jonni M. Kanerva (1989) Locative inversion in Chichewâ: a case study of factorization in grammar. *Linguistic Inquiry* 20: 1–50.

Bresnan, Joan, and Sam A. Mchombo (1987) Topic, pronoun, and agreement in Chichewâ. *Language* 63: 741–82.

Bresnan, Joan, and Sam A. Mchombo (1995) The Lexical Integrity Principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13: 181–254.

Brinton, Laurel J., and Elizabeth Closs Traugott (2005) *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

Brisard, Frank (ed.) (2002) *Grounding: The Epistemic Footing of Deixis and Reference*. Berlin: Mouton de Gruyter.

Bromhead, Helen (2009) *The Reign of Truth and Faith: Epistemic Expressions in 16th and 17th Century English*. Berlin: Mouton de Gruyter.

Brouland, Josephine (1855) *Langage mimique: spécimen d'un dictionnaire des signes*. Washington, DC: Gallaudet Archives.

Browman, Catherine P., and Louis M. Goldstein (1992) Articulatory phonology: an overview. *Phonetica* 49: 155–80.

Brown, Gillian, and George Yule (1983) *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, Penelope (2001) Learning to talk about motion UP and DOWN in Tzeltal: is there a language-specific bias for verb learning? In M. Bowerman and S. C. Levinson (eds.), *Language Acquisition and Conceptual Development*, pp. 512–43. Cambridge: Cambridge University Press.

Brown, Penelope, and Stephen C. Levinson (1987) *Politeness: Some Universals of Language Usage*. Cambridge: Cambridge University Press.

Brugè, Laura (2002) The position of demonstratives in the Extended Nominal Projection. In G. Cinque (ed.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures*, vol. 1, pp. 15–53. New York: Oxford University Press.

Bruner, Jerome S., Jacqueline J. Goodnow, and George A. Austin (1956) *A Study of Thinking*. New York: Wiley.

Bryant, John (2008) Best-fit Constructional Analysis. Doctoral dissertation, University of California, Berkeley.

Bühler, Karl (1982[1934]) *Spachtheorie: Die Darstellungsfunktion der Sprache*. Stuttgart: Fischer.

Bultinck, B. (2005) *Numerous Meaning: The Meaning of English Cardinals and the Legacy of Paul Grice*. Oxford: Elsevier.

Burchardt, Aljoscha, Katrin Erk, Anette Frank, Andrea Kowalski, and Sebastian Padó (2006) The SALSA corpus: a German corpus resource for lexical semantics. *Proceedings of LREC 2006, Workshop on Semantically Annotated Corpora* (Genoa), pp. 969–74. Paris: ELRA.

Burling, Robbins (2004) *The Language of the Modhupur Mandi (Garo)*, vol. 1: *Grammar*. New Delhi/Morganville, NJ: Bibliophile South Asia/Promilla.

Burton-Robert, Noel (ed.) (2007) *Pragmatics*. Basingstoke: Palgrave Macmillan.

Bush, Nathan (2001) Frequency effects and word-boundary palatalization in English. In Bybee and Hopper (2001: 255–80).

Butler, Christopher S. (1998) Collocational frameworks in Spanish. *International Journal of Corpus Linguistics* 3: 1–32.

Butler, Christopher S. (2003a) *Structure and Function: A Guide to Three Major Structural-Functional Theories*. Amsterdam: Benjamins.

Butler, Christopher S. (2003b) *Approaches to the Simple Clause*. Part 1 of Butler (2003a).

Butler, Christopher S. (2004) Notes towards an incremental implementation of the Role and Reference Grammar semantics-to-syntax mapping rules for English. In M. Hannay and G. Steen (eds.), *The English Clause: Usage and Structure*. Amsterdam: Benjamins, pp. 275–307.

Butler, Christopher S. (2008) Interpersonal meaning in the noun phrase. In Jan Rijkhoff and Daniel García Velasco (eds.), *The Noun Phrase in Functional Discourse Grammar*. Berlin: Mouton de Gruyter, pp. 221–61.

Butler, Christopher S., and Miriam Taverniers (2008) Layering in structural-functional grammars. *Linguistics* 46: 689–756.

Butt, David (1988) Randomness, order and the latent patterning of text. In R. Fawcett and D. Young (eds.), *Functions of Style*, pp. 74–97. London: Pinter.

Butt, Miriam (1995) *The Structure of Complex Predicates in Urdu*. Stanford, CA: CSLI Publications.

Butt, Miriam, Mary Dalrymple, and Anette Frank (1997) An architecture for linking theory in LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG97 Conference*. Stanford, CA: CSLI Publications.

Butt, Miriam, Mary Dalrymple, and Tracy Holloway King (eds.) (2006) *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan*. Stanford, CA: CSLI Publications.

Butt, Miriam, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi, and Christian Rohrer (2002) The parallel grammar project. *Proceedings of COLING-2002 Workshop on Grammar Engineering and Evaluation*, pp. 1–7.

Butt, Miriam, and Tracy Holloway King (eds.) (1996) *Proceedings of the LFG96 Conference*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (1998a) Interfacing Phonology with LFG. In Butt and King (1998b).

Butt, Miriam, and Tracy Holloway King (eds.) (1998b) *Proceedings of the LFG98 Conference*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (eds.) (2000a) *Argument Realization*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (2000b) Null elements in discourse structure. In K. V. Subbarao (ed.), *Papers from the NULLS Seminar*. Delhi: Motilal Banarasidas.

Butt, Miriam, and Tracy Holloway King (eds.) (2001a) *Proceedings of the LFG01 Conference*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (eds.) (2001b) *Time Over Matter: Diachronic Perspecitves on Morphosyntax*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (2004) The status of case. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, pp. 153–98. Dordrecht: Kluwer.

Butt, Miriam, and Tracy Holloway King (eds.) (2006) *Lexical Semantics in LFG*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (eds.) (2007) *Proceedings of the LFG07 Conference*. Stanford, CA: CSLI Publications.

Butt, Miriam, and Tracy Holloway King (eds.) (2009) *Proceedings of the LFG07 Conference*. Stanford, CA: CSLI Publications.

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño, and Frédérique Segond (1999) *A Grammar Writer's Cookbook*. Stanford, CA: CSLI Publications.

Butt, Miriam, María-Eugenia Niño, and Frédérique Segond (1996) Multilingual processing of auxiliaries within LFG. In Dafydd Gibbon (ed.), *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference*, pp. 111–22. Berlin: Mouton de Gruyter.

Bybee, Joan L. (1985a) *Morphology: a Study of the Relation Between Meaning and Form*. Amsterdam: Benjamins.

Bybee, Joan L. (1985b) Diagrammatic iconicity in stem–inflection relations. In John Haiman (ed.) *Iconicity in Syntax*, pp. 11–47. Amsterdam: Benjamins.

Bybee, Joan L. (1988a) The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining Language Universals*, pp. 350–79. Oxford: Blackwell.

Bybee, Joan L. (1988b) Morphology as lexical organization. In Michael Hammond and Michael Noonan (eds.), *Theoretical Morphology*, pp. 119–41. San Diego, CA: Academic Press.

Bybee, Joan L. (1988c) Semantic substance vs. contrast in the development of grammatical meaning. *Berkeley Linguistics Society* 14: 247–64.

Bybee, Joan L. (1995) Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–55.

Bybee, Joan L. (1998) The emergent lexicon. *Chicago Linguistic Society* 34, *The Panels*: 421–35.

Bybee, Joan L. (2000a) Lexicalization of sound change and alternating environments. In Michael Broe and Janet Pierrehumbert (eds.), *Papers in Laboratory Phonology 5: Acquisition and the Lexicon*, pp. 250–68. Cambridge: Cambridge University Press.

Bybee, Joan L. (2000b) The phonology of the lexicon: evidence from lexical diffusion. In Barlow and Kemmer (2000: 65–85).

Bybee, Joan L. (2001a) *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, Joan L. (2001b) Frequency effects on French liaison. In Bybee and Hopper (2001: 361–82).

Bybee, Joan L. (2002a) Sequentiality as the basis of constituent structure. In T. Givón and Bertram F. Malle (eds.), *The Emergence of Language out of Pre-language*, pp. 107–32. Amsterdam: Benjamins.

Bybee, Joan L. (2002b) Word frequency and context of use in the lexical diffusion of phonetically-conditioned sound change. *Language Variation and Change* 14: 261–90.

Bybee, Joan L. (2002c) Phonological evidence for examplar storage of multiword sequences. *Studies in Second Language Acquisition* 24: 215–21.

Bybee, Joan L. (2003a) Cognitive processes in grammaticalization. In Michael Tomasello (ed.), *The New Psychology of Language*, vol. 2, pp. 145–67. Mahwah, NJ: Erlbaum.

Bybee, Joan L. (2003b) Mechanisms of change in grammaticalization: the role of frequency. In Joseph and Janda (2003: 602–23).

Bybee, Joan L. (2006a) From usage to grammar: the mind's response to repetition. *Language* 82: 4: 711–33.

Bybee, Joan L. (2006b) Language change and universals. In Ricardo Mairal and Juana Gil (eds.), *Linguistic Universals*, pp. 179–94. Cambridge: Cambridge University Press.

Bybee, Joan L. (2006c) *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

Bybee, Joan L. (2008) Formal universals as emergent phenomena: the origins of Structure Preservation. In Jeff Good (ed.), *Language Universals and Language Change*, pp. 108–21. Oxford: Oxford University Press.

Bybee, Joan L., and David Eddington (2006) A usage-based approach to Spanish verbs of becoming. *Language* 82: 323–55.

Bybee, Joan L., John Haiman, and Sandra A. Thompson (eds.) (1997)*Essays on Language Function and Language Type.* Amsterdam: Benjamins.

Bybee, Joan L., and Paul J. Hopper (eds.) (2001) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.

Bybee, Joan L., and John McClelland (2005) Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *Linguistic Review* 22: 381–410.

Bybee, Joan L., and Carol Lynn Moder (1983) Morphological classes as natural categories. *Language* 59: 251–70.

Bybee, Joan L., William Pagliuca, and Revere D. Perkins (1991) Back to the future. In Traugott and Heine (1991b: 17–58).

Bybee, Joan L., Revere D. Perkins, and William Pagliuca (1994) *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Bybee, Joan L., and Joanne Scheibman (1999) The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37: 575–96.

Bybee, Joan L., and Dan I. Slobin (1982) Why small children cannot change language on their own: evidence from the English past tense. In A. Alqvist (ed.), *Papers from the Fifth International Conference on Historical Linguistics*, pp. 29–37. Amsterdam: Benjamins.

Bybee, Joan L., and Sandra A. Thompson. (1997) Three frequency effects in syntax. *Berkeley Linguistics Society* 23: 378–88.

Byrne, Richard W., and Andrew Whiten (eds.) (1988) *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Oxford University Press.

Caffarel, Alice (2004) The construal of a second-order semiosis in Camus' *L'Étranger*. In David Banks (ed.), *Text and Texture: Systemic Functional Viewpoints on the Nature and Structure of Text*, pp. 537–70. Paris: L'Harmattan.

Caffarel, Alice (2006a) *A Systemic Functional Grammar of French: From Grammar to Discourse*. London: Continuum.

Caffarel, Alice (2006b) Learning advanced French through SFL. In Heidi Byrnes (ed.), *Advanced Language Learning: The Contribution of Halliday and Vygotsky*, pp. 204–24. London: Continuum.

Caffarel, Alice, James R. Martin, and Christian M. I. M. Matthiessen (eds.) (2004) *Language Typology: A Functional Perspective*. Amsterdam: Benjamins.

Caffarel, Alice, and Elizabeth Rechniewski (2008) Ideological opposition in news stories: a French case study. In Elizabeth Thomson and Peter White (eds.), *Communicating Conflict: Multilingual Case Studies of the Rhetoric of the News Media*, pp. 25–49. London: Continuum.

Cahill, Aoife, Michael Burke, Ruth O'Donovan, Stefan Riezler, Josef van Genabith, and Andy Way (2008) Wide-coverage deep statistical parsing using automatic dependency structure annotation. *Computational Linguistics* 34: 81–124.

Cahill, Aoife, Martin Forst, Michael Burke, Mairéad McCarthy, Ruth O'Donovan, Christian Rohrer, Josef van Genabith, and Andy Way (2005) Treebank-based acquisition of multilingual unification grammar resources. *Research on Language and Computation* 3: 247–79.

Camazine, Scott, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz, and Eric Bonabeau (2001) *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

Čamdžić, Amela, and Richard Hudson (2007) Serbo-Croat clitics and Word Grammar. *Research in Language* (University of Lodz) 4: 5–50.

Campbell, Lyle (1991) Some grammaticalization changes in Estonian and their implications. In Traugott and Heine (1991a: 285–99).

Campbell, Lyle (2001) What's wrong with grammaticalization? *Language Sciences* 23: 113–61.

Campbell, Lyle, and Richard Janda (2001) Introduction: conceptions of grammaticalization and their problems. *Language Sciences* 23: 93–112.

Cantrall, William R. (1974) *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton.

Cara, Francesco, and Guy Politzer (1993) A comparison of conditional reasoning in English and Chinese. In J. Altarriba (ed.), *Cognition and Culture: A Cross-Cultural Approach to Cognitive Psychology*, pp. 283–97. Amsterdam: Elsevier Science.

Caramazza, Alfonso, and Bradford Z. Mahon (2006) The organization of conceptual knowledge in the brain: the future's past and some future directions. *Cognitive Neuropsychology* 23: 13–38.

Caramazza, Alfonso, and Kevin Shapiro (2004) Language categories in the brain. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, pp. 15–38. New York: Oxford University Press.

Cardinaletti, Anna (2004) Towards a cartography of syntactic positions. In L. Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures*, vol. 2, pp. 115–65. New York: Oxford University Press.

Carkin, Susan (2001) Pedagogic language in introductory classes: a multi-dimensional analysis of textbooks and lectures in biology and macroeconomics. PhD dissertation, Northern Arizona University.

Carlson, Greg N. (1983) Marking constituents. In F. Heny and B. Richards (eds.), *Linguistic Categories: Auxiliaries and Related Puzzles*, vol. 1, pp. 69–98. Dordrecht: Reidel.

Carpenter, Patricia, Akira Miyake, and Marcel Just (1994) Working memory constraints in comprehension: evidence from individual differences, aphasia, and aging. In M. Gernsbacher (ed.), *Handbook of Psycholinguistics*, pp. 1075–1122. San Diego, CA: Academic Press.

Carroll, John, and Michael Tanenhaus (1975) Prolegomena to a functional theory of word formation. *Chicago Linguistic Society Papers from the Parasession on Functionism*, pp. 47–62.

Carroll, John, and David Weir (2000) Encoding frequency information in stochastic parsing models. In H. Bunt and A. Nijholt (eds.), *Advances in Probabilistic Parsing and Other Parsing Technologies*, pp. 13–28. Dordrecht: Kluwer.

Carroll, Sean B. (2005) *Endless Forms Most Beautiful*. New York: Norton.

Carston, Robyn (1988) Implicature, explicature, and truth-theoretic semantics. In R. M. Kempson (ed.), *Mental Representations: The Interface between Language and Reality*, pp. 155–81. Cambridge: Cambridge University Press.

Carston, Robyn (1994) Metalinguistic negation and echoic use. *UCL Working Papers in Linguistics* 6: 321–39.

Carston, Robyn (1998) Postscript (1995) to Carston (1988). In A. Kasher (ed.), *Pragmatics: Critical Concepts*, vol. 4, pp. 464–79. London: Routledge.

Carston, Robyn (2001) Relevance theory and the saying/implicating distinction. *UCL Working Papers in Linguistics* 13: 1–34.

Carston, Robyn (2002) *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Carston, Robyn (2007) How many pragmatic systems are there? In M. J. Frápolli (ed.), *Saying, Meaning and Referring: Essays on François Recanati's Philosophy of Language*, pp. 18–48. Basingstoke: Palgrave Macmillan.

Carston, Robyn, and Eun-Ju Noh (1995) A truth-functional account of metalinguistic negation, with evidence from Korean. *UCL Working Papers in Linguistics* 7: 1–26.

Carston, Robyn, and George Powell (2005) Relevance theory: new directions and developments. *UCL Working Papers in Linguistics* 17: 279–99.

Carter, A. (1974) Communication in the sensory-motor period. PhD dissertation, University of California, Berkeley.

Casenhiser, Devin, and Adele E. Goldberg (2005) Fast mapping of a phrasal form and meaning. *Developmental Science* 8: 500–508.

Chafe, Wallace (ed.) (1980) *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.

CHAFE, WALLACE (1998) Language and the flow of thought. In Michael Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, pp. 93–111. Mahwah, NJ: Erlbaum.

CHAMBERS, J. K., and PETER TRUDGILL (1980) *Dialectology*. Cambridge: Cambridge University Press.

CHANG, NANCY, and EVA MOK (2006) Putting context in constructions. *The Fourth International Conference on Construction Grammar* (ICCG4, Tokyo).

CHAPMAN, SIOBHAN (2005) *Paul Grice: Philosopher and Linguist*. London: Palgrave Macmillan.

CHARNIAK, EUGENE (1997) Statistical techniques for natural language parsing. *AI Magazine*, Winter: 32–43.

CHEN, JENN-YEU (2007) Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition* 104: 427–36.

CHENG, LISA L.-S., and RINT SYBESMA (1999) Bare and not so bare nouns and the structure of NP. *Linguistic Inquiry* 304: 509–42.

CHERNIAK, CHRISTOPHER, ZEKERIA MOKHTARZADA, RAUL RODRIGUEZ-ESTEBAN, and MARK CHANGIZI (2004) Global optimization of cerebral cortex layout. *Proceedings of the National Academy of Sciences* 101: 1081–6.

CHESI, CRISTIANO (2005) Phases and cartography in linguistic computation: toward a cognitively motivated computational model of linguistic competence. PhD dissertation, University of Siena.

CHEYNEY, DOROTHY L., and ROBERT M. SEYFARTH (2007) *Baboon Metaphysics: The Evolution of a Social Mind*. Chicago: University of Chicago Press.

CHI, ZHIYI and STUART GEMAN (1998) Estimation of probabilistic context-free grammars. *Computational Linguistics* 24: 299–305.

CHIANG, DAVID (2003) Statistical parsing with an automatically extracted tree adjoining grammar. 2003. In R. Bod, R. Scha, and K. Sima'an (eds.), *Data Oriented Parsing*, pp. 299–316. Stanford, CA: CSLI Publications.

CHIERCHIA, GENNARO (1998) Reference to kinds across languages. *Natural Language Semantics* 6: 339–405.

CHIERCHIA, GENNARO (2004) Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (ed.) *Structures and Beyond*, pp. 39–103. Oxford: Oxford University Press.

CHOI, SOONJA, and MELISSA BOWERMAN (1991) Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition* 41: 81–121.

CHOMSKY, NOAM (1951) Morphophonemics of Modern Hebrew. MA thesis, University of Pennsylvania. Published 1979, New York: Garland.

CHOMSKY, NOAM (1957) *Syntactic Structures*. The Hague: Mouton.

CHOMSKY, NOAM (1955) The logical structure of linguistic theory. PhD thesis, University of Pennsylvania.

CHOMSKY, NOAM (1959) Review of B. F. Skinner's *Verbal Behavior*. *Language* 35: 26–58.

CHOMSKY, NOAM (1961) Some methodological remarks on generative grammar. *Word* 17: 219–39.

CHOMSKY, NOAM (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

CHOMSKY, NOAM (1966)*Cartesian Linguistics*. New York: Harper & Row.

CHOMSKY, NOAM (1968/1972) *Language and Mind*. New York: Harcourt, Brace & World.

Chomsky, Noam (1970) Remarks on nominalization. In Roderick A. Jacobs and Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, pp. 184–221. Waltham, MA: Ginn.

Chomsky, Noam (1972) *Studies on Semantics in Generative Grammar*. The Hague: Mouton.

Chomsky, Noam (1973) Conditions on transformations. In Stephen R. Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, pp. 232–86. New York: Holt, Rinehart, & Winston.

Chomsky, Noam (1977) On wh-movement. In Peter W. Culicover, Thomas Wasow, and Adrian Akmajian (eds.), *Formal Syntax*, pp. 71–132. New York: Academic Press.

Chomsky, Noam (1980) *Rules and Representations*. New York: Columbia University Press.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam (1982) *The Generative Enterprise*. Berlin: Mouton de Gruyter.

Chomsky, Noam (1986a) *Knowledge of Language: Its Nature, Origin, and Use*. Westport, CT: Praeger.

Chomsky, Noam (1986b) *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1988) *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.

Chomsky, Noam (1992) *A Minimalist Program for Linguistic Theory*. Cambridge, MA: MIT Press.

Chomsky, Noam (1993) A minimalist program for linguistic theory. In Kenneth Hale and Samuel J. Keyser (eds.), *The View from Building 20*, pp. 1–52. Cambridge, MA: MIT Press. Reprinted in Chomsky (1995: 167–217).

Chomksy, Noam (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000a) Minimalist inquiries: the framework. In Roger Martin, David Michaels, and Juan Uriagereka (eds.), *Step by Step*, pp. 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam (2000b) *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Chomsky, Noam (2001a) *Beyond Explanatory Adequacy*. MIT Occasional Papers in Linguistics 20. Cambridge, MA: MIT Dept of Linguistics.

Chomsky, Noam (2001b) Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, pp. 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam (2004) Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond*, pp. 104–31. Oxford: Oxford University Press.

Chomsky, Noam (2005) Three factors in language design. *Linguistic Inquiry* 36: 1–22.

Chomsky, Noam (2007) Approaching UG from below. In Uli Sauerland and Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, pp. 1–30. Berlin: Mouton de Gruyter.

Chomsky, Noam (2008) On phases. In Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory*, pp. 133–66. Cambridge, MA: MIT Press.

Chomsky, Noam (2009) Opening remarks. In M. Piattelli-Palmarini, P. Salaburu, and J. Uriagereka (eds.), *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*, 13–42. Oxford: Oxford University Press.

Chomsky, Noam and Morris Halle (1968) *The Sound Pattern of English*. New York: Harper & Row.

Chomsky, Noam, and Howard Lasnik (1977) Filters and control. *Linguistic Inquiry* 28: 425–504.

Chomsky, Noam, and Howard Lasnik (1993) The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, *Syntax: An International Handbook of Contemporary Research*, pp. 506–69. Berlin: de Gruyter.

Chomsky, Noam, and George A. Miller (1963) Introduction to the formal analysis of natural languages. In R. Luce, R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, pp. 269–321. New York: Wiley.

Christie, Frances, and J. R. Martin (2006) *Language, Knowledge and Pedagogy: Functional Linguistic and Sociological Perspectives*. London: Continuum.

Chung, Sandra (1998) *The Design of Agreement*. Chicago: University of Chicago Press.

Chung, Sandra, and James McCloskey (1987) Government, barriers and small clauses in Modern Irish. *Linguistic Inquiry* 18: 173–237.

Church, Alonzo (1940) A formulation of the simple theory of types. *Journal of Symbolic Logic* 5: 56–68.

Church, Andrea Brandilyn (2005) Linguistic relativity, nonlinguistic influences, and spatial conceptualization: ascertaining the relationship between language and environment. Honors thesis, Univeristy of Oregon, Eugene.

Cinque, Guglielmo (1990) *Types of A-bar Dependencies*. Cambridge, MA: MIT Press.

Cinque, Guglielmo (1994) On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (eds.), *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*, pp. 85–110. Washington, DC: Georgetown University Press.

Cinque, Guglielmo (1999) *Adverbs and Functional Heads: A Cross-linguistic Perspective*. New York: Oxford University Press.

Cinque, Guglielmo (ed.) (2002) *Functional Structure in DP and IP: The Cartography of Syntactic Structures*, vol. 1. New York: Oxford University Press.

Cinque, Guglielmo (2004) Issues in adverbial syntax. *Lingua* 114: 683–710.

Cinque, Guglielmo (2005) Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36: 315–32.

Cinque, Guglielmo (2006a) *Restructuring and Functional Heads: The Cartography of Syntactic Structures*, vol. 4. New York: Oxford University Press.

Cinque, Guglielmo (2006b) Are all languages 'numeral classifier languages'? *Rivista di grammatica generativa* 31: 119–22.

Cinque, Guglielmo (2009) The fundamental left–right asymmetry of natural languages. In S. Scalise, E. Magni, and A. Bisetto (eds.), *Universals of Language Today*, pp. 165–84. Dordrecht: Springer.

Cinque, Guglielmo (forthcoming) *The Syntax of Adjectives: A Comparative Study*. Cambridge, MA: MIT Press.

Cinque, Guglielmo, and Luigi Rizzi (forthcoming) *Mapping Spatial PPs: The Cartography of Syntactic Structures*, vol. 6. New York: Oxford University Press.

Clark, Alexander (2000) Inducing syntactic categories by context distribution clustering. *Proceedings CoNLL 2000*, pp. 91–4.

Clark, Alexander (2001) Unsupervised induction of stochastic context-free grammars using distributional clustering. *Proceedings CoNLL 2001*, pp. 105–112.

Clark, Alexander, and Remy Eyraud (2006) Learning auxiliary fronting with grammatical inference. *Proceedings CoNLL 2006* (New York).

CLARK, BILLY (1993) Relevance and 'pseudo-imperatives'. *Linguistics and Philosophy* 16: 79–121.

CLARK, BRADY Z. (2004) A stochastic Optimality Theory approach to syntactic change. PhD thesis, Stanford University.

CLARK, BRADY Z. (2005) On stochastic grammar. *Language* 81: 207–17.

CLARK, EVE (1990) On the pragmatics of contrast. *Journal of Child Language* 17: 417–31.

CLARK, EVE (1992) Conventionality and contrast: pragmatic principles with lexical consequences. In Adrienne Lehrer and E. F. Kittay (eds.), *Frames, Fields and Contrasts*, pp. 171–88. Hillsdale, NJ: Erlbaum.

CLARK, EVE (1993) *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.

CLEMENTS, G. N. (1990) The role of the sonority cycle in core syllabification. In John Kingston and Mary Beckman (eds.), *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, pp. 283–333. New York: Cambridge University Press.

COENE, ANN (2006) *Lexikalische Bedeutung, Valenz und Koerzion*. Germanistische Linguistik Monographien 19. Hildesheim: Olms.

COETZEE, ANDRIES (2001) Nasal-only syllables in Tswana. In Andriew Coetzee, René Kager, and Joe Pater (eds.), *Ruling Papers II: Relations Between Phonological Models And Experimental Data*. Special issue of *Phonology*.

COETZEE, ANDRIES, and JOE PATER (forthcoming) The place of variation in phonological theory. In John Goldsmith, Jason Riggle, and Alan Yu (eds.), *The Handbook of Phonological Theory*, 2nd edn. Malden, MA: Blackwell.

COHEN, ANTHONY (1987) Morfologie op heterdaad betrapt. *Forum der Letteren* 28: 90–97.

COJOCARU, DANA (2003) *Romanian Grammar*. Durham, NC: Slavic and East European Language Research Center, Duke University.

COLE, PETER and MORGAN, JERRY (eds.) (1975) *Speech Acts*, New York: Academic Press.

COLEMAN, JOHN, and JANET PIERREHUMBERT (1997) Stochastic phonological grammars and acceptability. In *Computational Phonology: Third Meeting of the Association for Computational Linguistics Special Interest Group in Computational Phonology*, pp. 49–56.

COLLEMAN, TIMOTHY, BART DEFRANCQ, FILIP DEVOS, and DIRK NOËL (2004) *The Contragram Dutch–French–English Contrastive Verb Valency Dictionary*. Ghent: University of Ghent (http://www.contragram.ugent.be/cvvd.htm)

Collins Dictionary of the English Language (1985) Glasgow: Collins.

COLLINS, MICHAEL (1996) A new statistical parser based on bigram lexical dependencies. *Proceedings of the Association for Computational Linguistics* 34: 184–91.

COLLINS, MICHAEL (1999) Head-driven statistical models for natural language processing. PhD thesis, University of Pennsylvania.

COLLINS, MICHAEL, and NIGEL DUFFY (2001) Convolution kernels for natural language. *Neural Information Processing Systems 2001*.

COLLINS, PETER (1991) *Cleft and Pseudo-cleft Constructions in English*. London: Routledge.

COLLINS, PETER (1995) The indirect object construction in English: an informational approach. *Linguistics* 33: 35–49.

COMRIE, BERNARD (1981) *Language Universals and Linguistic Typology*. Oxford: Blackwell.

CONNOR, ULLA (2004) The genre of grant proposals: a corpus linguistic analysis. In Connor and Upton (2003: 235–56).

CONNOR, ULLA, and THOMAS UPTON (2003) Linguistic dimensions of direct mail letters. In C. Meyer and P. Leistyna (eds.), *Corpus Analysis: Language Structure and Language Use*, pp. 71–86. Amsterdam: Rodopi.

Connor-Linton, Jeff (2001) Authors style and world-view: a comparison of texts about nuclear arms policy. In Conrad and Biber (2001: 84–93).

Connor-Linton, Jeff, and E. Shohamy (2001) Register variation, oral proficiency sampling, and the promise of multi-dimensional analysis. In Conrad and Biber (2001: 124–37).

Conrad, Susan (1996a) Academic discourse in two disciplines: professional writing and student development in biology and history. PhD dissertation, Northern Arizona University.

Conrad, Susan (1996b) Investigating academic texts with corpus-based techniques: an example from biology. *Linguistics and Education* 8: 299–326.

Conrad, Susan (2001a) Variation among disciplinary texts: a comparison of textbooks and journal articles in biology and history. In Conrad and Biber (2001: 94–107).

Conrad, Susan (ed.) (2001b) *Variation in English: Multi-dimensional Studies*. London: Longman.

Conrad, Susan, and Douglas Biber (eds.) (2001) *Multi-dimensional Studies of Register Variation in English*. London: Longman.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz, and Clin Phillips (2008) Equal treatment for all antecedents: how children succeed with Principle B. MS, Dept of Linguistics, University of Maryland.

Copestake, Ann, and Ted Briscoe (1995) Semi-productive polysemy and sense extension. *Journal of Semantics* 12: 15–67.

Corbett, Greville G. (2000) *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G., and Ian R. L. Davies (1997) Establishing basic color terms: measures and techniques. In C. L. Hardin and L. Maffi (eds.), *Color Categories in Thought and Language*, pp. 197–223. New York: Cambridge University Press.

Cormons, Boris (1999) Analyse et désambiguisation: une approche à base de corpus (data-oriented parsing) pour les représentations lexicales fonctionelles. PhD thesis, University of Rennes.

Cortes, Viviana (2002) Lexical bundles in academic writing in history and biology. Doctoral dissertation, Northern Arizona University.

Cortes, Viviana (2004) Lexical bundles in published and student disciplinary writing: examples from history and biology. *English for Specific Purposes* 23: 397–423.

Coulson, Seanna, Jonathon King, and Marta Kutas (1998‘) Expect the unexpected: event-related brain response to morphosyntactic violations. *Language and Cognitive Processes* 13: 21–58.

Craig, Colette (ed.) (1986) *Noun Classes and Categorization*. Amsterdam: Benjamins.

Crain, Stephen (1991) Language acquisition in the absence of experience. *Behavorial and Brain Sciences* 14: 597–612.

Crain, Stephen, and Mineharu Nakayama (1987) Structure dependence in grammar formation. *Language* 63: 522–43.

Crain, Stephen, and P. Pietroski (2002) Why language acquisition is a snap. *Linguistic Review* 19: 163–83.

Creider, Chet, and Richard Hudson (2006) Case agreement in Ancient Greek: implications for a theory of covert elements. In Kensei Sugayama and Richard Hudson (eds.), *Word Grammar: New Perspectives on a Theory of Language Structure*, pp. 35–53. London: Continuum.

Crick, Francis (1998) *What Mad Pursuit: A Personal View of Scientific Discovery*. New York: Basic Books.

Cristofaro, Sonia (2008) Grammatical categories and relations: universality vs. language-specificity and construction-specificity. MS, University of Pavia.

Crocker, Matthew, and Frank Keller (2006) Probabilistic grammars as models of gradience in language processing. In G. Fanselow, C. Féry, R. Vogel, and M. Schlesewsky (eds.), *Gradience in Grammar: Generative Perspectives*. Oxford: Oxford University Press, pp. 227–45.

Croft, William (2000) *Explaining Language Change: An Evolutionary Approach*. London: Longman.

Croft, William (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William (2002) *Typology and Universals*. Cambridge: Cambridge University Press.

Croft, William (2003) *Typology and Universals*, 2nd edn. Cambridge: Cambridge University Press.

Croft, William (2006) Typology. In Aronoff and Rees-Miller (2006: 337–68).

Croft, William (2007a) Construction grammar. In D. Geeraerts and H. Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, pp. 463–508. Oxford: Oxford University Press.

Croft, William (2007b) The origins of grammar in the verbalization of experience. *Cognitive Linguistics* 18: 339–82.

Croft, William, and D. A. Cruse (2002) *Cognitive Grammar*. Cambridge: Cambridge University Press.

Cross, Marilyn (1993) Collocation in computer modeling of lexis as most delicate grammar. In Mohsen Ghadessy (ed.), *Register Analysis: Theory and Practice*, pp. 196–220. London: Pinter.

Crossley, Scott A., and Max M. Louwerse (2007) Multi-dimensional register classification using bigrams. *International Journal of Corpus Linguistics* 12: 453–78.

Crosswhite, Katherine (1999) Vowel reduction in Optimality Theory. Doctoral dissertation, University of California, Los Angeles.

Crosswhite, Katherine (2001) Predicting syllabicity and moraicity in Dihovo Macedonian. In J. S. Magnuson and Katherine Crosswhite (eds.), *University of Rochester Working Papers in the Language Sciences*, vol. 2, pp. 1–22.

Crouch, Richard (2005) Packed rewriting for mapping semantics to KR. *Sixth International Workshop on Computational Semantics (IWCS-6)* (Tilburg), pp. 103–114.

Crouch, Richard (2006) Packed rewriting for mapping text to semantics and KR. In Butt et al. (2006: 389–416).

Crouch, Richard, Mary Dalrymple, Ron Kaplan, Tracy King, John Maxwell, and Paula Newman (2008) *XLE Documentation*. Palo Alto, CA: Palo Alto Research Center.

Crouch, Richard, and Tracy Holloway King (2006) Semantics via f-structure rewriting. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG06 Conference*, pp. 145–165. Stanford, CA: CSLI Publications.

Cruschina, Silvio (2006) Informational focus in Sicilian and the left periphery. In M. Frascarelli (ed.), *Phases of Interpretation*, pp. 363–85. Berlin: Mouton de Gruyter.

Cruschina, Silvio (2008) Discourse-related features and the syntax of peripheral positions: a comparative study of Sicilian and other Romance languages. PhD dissertation, University of Cambridge.

Cruse, D. Alan (1986) *Lexical Semantics*. Cambridge: Cambridge University Press.

Csomay, Eniko (2005) Linguistic variation within university classroom talk: a corpus-based perspective. *Linguistics and Education* 15: 243–74.

Culicover, Peter W. (1999) *Syntactic Nuts*. Oxford: Oxford University Press.

Culicover, Peter W., and Ray Jackendoff (1999) The view from the periphery. *Linguistic Inquiry* 40: 543–71.

Culicover, Peter W., and Ray Jackendoff (2005) *Simpler Syntax*. Oxford: Oxford University Press.

Culicover, Peter W., and Andrzej Novak (2003) *Dynamical Grammar*. Oxford: Oxford University Press.

Cysouw, Michael, Mihai Albu, and Andreas Dress (2007) Analyzing feature consistency using dissimilarity matrices. *Strachtypologie & Universalienforschung*, 61(3): 263–279.

Dąbrowska, Eva, and Elena Lieven (2005) Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16: 437–74.

Daelemans, Walter, and Antal van den Bosch (2005) *Memory-Based Language Processing*. Cambridge: Cambridge University Press.

Dalrymple, Mary (1993) *The Syntax of Anaphoric Binding*. Stanford, CA: CSLI Publications.

Dalrymple, Mary (ed.) (1999) *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. Cambridge, MA: MIT Press.

Dalrymple, Mary (2001) *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Dalrymple, Mary, and Ronald M. Kaplan (2000) Feature indeterminacy and feature resolution. *Language* 74: 759–98.

Dalrymple, Mary, Ronald M. Kaplan, and Tracy Holloway King (2001) Weak crossover and the absence of traces. In Butt and King (2001a: 66–82).

Dalrymple, Mary, Ronald M. Kaplan, and Tracy Holloway King (2007) The absence of traces: evidence from weak crossover. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling, and Chris Manning (eds.), *Architectures, Rules, and Preferences: Variations on Themes by Joan W. Bresnan*, p. 27. Stanford, CA: CSLI Publications.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen (eds.) (1995) *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Dalrymple, Mary, John Lamping, and Vijay Saraswat (1993) LFG semantics via constraints. *Proceedings of the Sixth Meeting of the European ACL*, pp. 97–105. European Chapter of the Association for Computational Linguistics, University of Utrecht.

Darnell, Michael, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley (eds.) (1999) *Functionalism and Formalism in Linguistics*, vol. 1: *General Papers*. Amsterdam: Benjamins.

Darwin, Charles (1859[1964]) *The Origin of Species*. Cambridge, MA: Harvard University Press.

Davidson, David (1967) The logical form of action sentences. In Nicholas Rescher (ed.), *The Logic of Decision and Action*, pp. 81–94. Pittsburgh: University of Pittsburgh Press.

Davies, Ian R. L. (1998) A study of colour grouping in three languages: a test of the linguistic relativity hypothesis. *British Journal of Psychology* 89: 433–52.

Davies, Ian R. L., and Greville G. Corbett (1997) A cross-cultural study of colour grouping: evidence for weak linguistic relativity. *British Journal of Psychology* 88: 493–517.

Davies, Ian R. L., Paul T. Sowden, David T. Jerrett, Tiny Jerrett and Greville G. Corbett (1998) A cross-cultural study of English and Setswana speakers on a colour triads task: a test of the Sapir-Whorf hypothesis. *British Journal of Psychology* 89: 1–15.

Davies, William D., and Stanley Dubinsky (eds.) (2004) *The Grammar of Raising and Control: A Course in Syntactic Argumentation*. Oxford: Blackwell.

Davis, Wayne A. (1998) *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*. Cambridge: Cambridge University Press.

Davis, Wayne A. (2007) How normative is implicature? *Journal of Pragmatics* 39: 1655–72.

Dawkins, Richard (1976) *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, Richard (1986) *The Blind Watchmaker*. New York: Norton.

Dayal, Veneeta (1999) Bare NPs, reference to kinds and incorporation. *Proceedings of SALT* 9. Ithaca, NY: Cornell University Press.

Dayal, Veneeta (2004) Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27: 393–450.

de Jorio, Andrea (2000[1832]) *Gesture in Naples and Gesture in Classical Antiquity*, a translation of *La mimica degli antichi investigata nel gestire napoletano* [*Gestural Expression of the Ancients in the Light of Neapolitan Gesturing*]. Bloomington: Indiana University Press.

de Lacy, Paul (2002a) The formal expression of markedness. Doctoral dissertation, University of Massachusetts, Amherst.

de Lacy, Paul (2002b) The interaction of tone and stress in Optimality Theory. *Phonology* 19: 1–32.

de Lacy, Paul (2004) Markedness conflation in Optimality Theory. *Phonology* 21: 145–99.

de Lacy, Paul (2007) Freedom, interpretability, and the Loop. In Sylvia Blaho, Patrick Bye, and Martin Krämer (eds.), *Freedom of Analysis?*, pp. 86–118. Berlin: Mouton de Gruyter.

de Villiers, Jill G., and Peter A. de Villiers (2000) Linguistic determinism and the understanding of false beliefs. In P. Mitchell and K. J. Riggs (eds.), *Children's Reasoning and the Mind*, pp. 191–228. Hove: Psychology Press.

de Villiers, Jill G., and Peter A. de Villiers (2003) Language for thought: coming to understand false beliefs. In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind*, pp. 335–84. Cambridge, MA: MIT Press.

de Waal, Frans (2001) *The Ape and the Sushi Master*. New York: Basic Books.

Deacon, Terrence (1997) *The Symbolic Species*. New York: Norton.

Deane, Paul (1992) *Grammar in Mind and Brain: Explorations in Cognitive Syntax*. New York: Mouton de Gruyter.

Decety, Jean, and Philip L. Jackson (2006) A social neuroscience perspective on empathy. *Current Directions in Psychological Science* 15: 54–8.

Decety, Jean, and Jessica A. Sommerville (2003) Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences* 7: 527–33.

Delfitto, Denis, and Jan Schroten (1991) Bare plurals and the number affix in DP. *Probus* 3: 155–85.

Dell, François, and Mohamed Elmedlaoui (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7: 105–30.

Demestre, Josep, Sheila Meltzer, José García-Albea, and Andreu Vigil (1999) Identifying the null aubject: evidence from event-related brain potentials. *Journal of Psycholinguistic Research* 28: 293–312.

Dempster, A., N. Laird, and D. Rubin (1977) Maximum likelihood from incomplete data via the EM Algorithm. *Journal of the Royal Statistical Society* 39: 1–38.

Dennett, Daniel C. (1995) *Darwin's Dangerous Idea*. New York: Simon & Schuster.
Dennis, Simon (2005) An exemplar-based approach to unsupervised parsing. *Proceedings of the Twenty Seventh Conference of the Cognitive Science Society*.
di Cristo, Albert (1998) Intonation in French. In Daniel Hirst and Albert Di Cristo (eds.), *Intonation Systems: A Survey of Twenty Languages*, pp. 195–218. Cambridge: Cambridge University Press.
Di Sciullo, Anna-Maria, and Edwin Williams (1987) *On the Definition of Word*. Cambridge, MA: MIT Press.
Dial, Kenneth E., Brandon D. Jackson, and Paolo Segre (2008) A fundamental avian wing-stroke provides new perspective on the evolution of flight. *Nature* 451: 985–89.
Dickens, Patrick. (1992) Ju/'hoan grammar. TS, Johannesburg.
Dickens, Patrick (2005) *A Concise Grammar of Ju/'hoan with a Ju/'hoan-English Glossary and a Subject Index*, ed. Rainer Vossen and Megan Biesele. Cologne: Köppe.
Diessel, Holger (1999) *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam: Benjamins.
Diessel, Holger (2005) *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.
Diewald, Gabriele (1997) *Grammatikalisierung: eine Einführung in Sein und Werden grammatischer Formen*. Tübingen: Niemeyer.
Dik, Simon C. (1978) *Functional Grammar.* Amsterdam: North-Holland.
Dik, Simon C. (1986) On the notion 'functional explanation'. *Belgian Journal of Linguistics* 1: 11–52.
Dik, Simon C. (1997)*The Theory of Functional Grammar*, vols. 1 and 2. Berlin: Mouton de Gruyter.
Dimmendaal, Gerrit J. (2000) Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42: 214–61.
Dixon, R. M. W. (1972) *The Dyirbal language of North Queensland.* Cambridge: Cambridge University Press.
Dixon, R. M. W. (1982) *Where Have All the Adjectives Gone? And Other Essays in Semantics and Syntax*. Berlin: Mouton.
Dixon, R. M. W. (1997) *The Rise and Fall of Languages.* Cambridge: Cambridge University Press.
Dixon, R. M. W. (2004) Adjective classes in typological perspective. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Adjective Classes: A Cross-linguistic Typology*, pp. 1–49. Oxford: Oxford University Press.
Dixon, R. M. W., and Alexandra Y. Aikhenvald (2002) Word: a typological framework. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Word: A Cross-linguistic Typology*, pp. 1–41. Cambridge: Cambridge University Press.
Dobzhansky, Theodosius (1973) Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* 35: 125–9.
Dodge, Ellen (2009) Conceptual and constructional composition. Doctoral dissertation, University of California, Berkeley.
Dodge, Ellen, and George Lakoff (2005) Image schemas: from linguistic analysis to neural grounding. In Beate Hampe (ed.), *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, pp. 57–91. Berlin: Mouton de Gruyter.
Dodwell, Charles R. (2000) *Anglo-Saxon Gestures and the Roman Stage*. Cambridge: Cambridge University Press.

DONNELLAN, KEITH S. (1966) Reference and definite descriptions. *Philosophical Review* 75: 281–304.

DORON, EDITH (2003) Bare singular reference to kinds. *Proceedings of SALT* 13, Cornell Working Papers in Linguistics, Ithaca, NY.

DOWTY, DAVID R. (1979) *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.

DOWTY, DAVID R. (1991) Thematic proto-roles and argument selection. *Language* 67: 547–619.

DOWTY, DAVID R. (2007) Compositionality as an empirical problem. In C. Barker and P. Jacobson (eds.), *Direct Compositionality*, pp. 23–101. Oxford: Oxford University Press.

DRACH, ERICH (1937) *Grundgedanken der deutschen Satzlehre.* Frankfurt am Main. (Usually cited from 4th edn, Darmstadt: Wissenschaftliche Buchgesellschaft, 1963.)

DRACHMAN, GABERELL (1978) Child language and language change: a conjecture and some refutations. In Jaček Fisiak (ed.), *Recent Developments in Historical Phonology*, pp. 123–44. The Hague: Mouton.

DRYER, MATTHEW S. (1989) Article–noun order. *Chicago Linguistic Society* 25: 83–97.

DRYER, MATTHEW S. (1997) Are grammatical relations universal? In Bybee et al. (1997: 115–43).

DRYER, MATTHEW S. (1998) Why statistical universals are better than absolute universals. *Chicago Linguistics Society* 33: 123–45.

DRYER, MATTHEW S. (2005a) Order of subordinator and adverbial subordinate clause. In *WALS*, pp. 382–5.

DRYER, MATTHEW S. (2005b) Order of subject, object, and verb. In In *WALS*, pp. 386–9.

DRYER, MATTHEW S. (2005c) Position of interrogative phrases in content questions. In *WALS*, pp. 378–81.

DRYER, MATTHEW S. (2005d) Definite articles. In *WALS*, pp. 154–7.

DRYER, MATTHEW S. (2005e) Indefinite articles. In *WALS*, pp. 158–61.

DRYER, MATTHEW S. (2006a) Functionalism and the theory-metalanguage confusion. In Wiebe et al. (2006: 27–59).

DRYER, MATTHEW S. (2006b) Descriptive theories, explanatory theories, and basic linguistic theory. In Felix Ameka, Alan Dench, and Nicholas Evans (eds.), *Catching Language: The Standing Challenge of Grammar Writing*, pp. 207–34. Berlin: Mouton de Gruyter.

DRYER, MATTHEW S. (2007) Noun phrase structure. In Timothy Shopen (ed.), *Complex Constructions, Language Typology and Syntactic Description*, vol. 2, 2nd edn, pp. 151–205. Cambridge: Cambridge University Press.

DUBOIS, JOHN (1985) Competing motivations. In John Haiman (ed.), *Iconicity in Syntax*, pp. 343–65. Amsterdam: Benjamins.

DURIE, MARK (1985) *A Grammar of Acehnese*. Dordrecht: Foris.

DURIE, MARK (1987) Grammatical relations in Acehnese. *Studies in Language* 11: 365–99.

DURRELL, MARTIN (2002) *Hammer's German Grammar and Usage*, 4th edn. London: Arnold.

DURRLEMAN, STEPHANIE (2007) The syntax of Jamaican Creole: a cartographic perspective. PhD dissertation, University of Geneva.

EBERT, KAREN H. (1991) Vom Verbum dicendi zur Konjunktion: Ein Kapitel universaler Grammatikentwicklung. In Bisang and Rinderknecht (1991: 77–95).

EISENBERG, PETER (2006) *Grundriss der deutschen Grammatik*, vol. 2:*Der Satz*, 3rd edn. Stuttgart: Metzler.

Elenbaas, Nine, and René Kager (1999) Ternary rhythm and the Lapse constraint. *Phonology* 16: 273–330.

Ellis, Nick C. (2002) Frequency effects in language processing and acquisition. *Studies in Second Language Acquisition* 24: 143–88.

Ellis, Nick C., and Diane Larsen-Freeman (2006) Language emergence: implications for applied linguistics (introduction to the special issue). *Applied Linguistics* 27.4: 558–89.

Ellsworth, Michael, Kyoko Ohara, Carlos Subirats, and Thomas Schmidt (2006) Frame-semantic analysis of motion scenarios in English, German, Spanish, and Japanese. *Fourth International Conference on Construction Grammar* (Tokyo).

Elman, Jeff, and Elizabeth Bates (1997) Response to letters. *Science* 276: 1180.

Emanatian, Michele. (1992) Chagga 'come' and 'go': metaphor and the development of tense-aspect. *Studies in Language* 16: 1–33.

Embick, David, and Rolf Noyer (2007) Distributed morphology and the syntax–morphology interface. In Gillian Ramchand and Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, pp. 289–324. Oxford: Oxford University Press.

Emonds, Joseph (1970) *Root and Structure Preserving Transformations*. Bloomington, Indiana: Indiana University Linguistics Club.

Emonds, J. (1978) The verbal complex V′ — V in French. *Linguistic Inquiry* 9: 151–75.

Emons, Rudolf (1974) *Valenzen englischer Prädikatsverben*. Tübingen: Niemeyer.

Enç, Murvet (1991) The semantics of specificity. *Linguistic Inquiry* 22: 1–25.

Enfield, Nick J. (2000) On linguocentrism. In M. Pütz and M. Verspoor (eds.), *Explorations in Linguistic Relativity*, pp. 125–57. Amsterdam: Benjamins.

Endo, Yoshio (2007) *Locality and Information Structure: A Cartographic Approach to Japanese*. Amsterdam: Benjamins.

Engel, Ulrich (1988) *Deutsche Grammatik: Vollständige Darstellung der deutschen Gegenwartssprache.* Heidelberg: Groos.

Engel, Ulrich (1994) *Syntax der deutschen Gegenwartssprache*. Berlin: Schmidt.

Engel, Ulrich (1999) *Deutsch-polnische kontrastive Grammatik* (2 vols). Heidelberg: Groos.

Engel, Ulrich (2004) *Deutsche Grammatik: Neubearbeitung*. Munich: Iudicium.

Engel, Ulrich, Mikai Isbasescu, Speranţa Stănescu, and Octavian Nicolae (1993) *Kontrastive Grammatik deutsch-rumänisch*. Heidelberg: Groos.

Engel, Ulrich and Pavica Mrazović (eds.) (1986) *Kontrastive Grammatik deutsch-serbokroatisch* (2 vols). Munich: Sagner.

Engelen, Bernhard (2003) Die Wortartenlehre bei Lucien Tesnière. In Ágel et al. (2003: 100–108).

Epstein, Samuel D., Erich Groat, Ruriko Kawashima, and Hisatsugu Kitahara (1998) *The Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.

Ericsson, K. Anders, and Walter Kintsch (1995) Long term working memory. *Psychological Review* 102: 211–45.

Erman, Britt, and Beatrice Warren (2000) The idiom principle and the open choice principle. *Text* 20: 29–62.

Eroms, Hans-Werner (2000) *Syntax der deutschen Sprache*. Berlin: de Gruyter.

Eroms, Hans-Werner, and Hans Jürgen Heringer (2003) Dependenz und lineare Ordnung. In Ágel et al. (2003: 247–63).

Ervin-Tripp, Susan (1972) On sociolinguistic rules: alternation and co-occurrence. In J. J. Gumperz and D. Hymes (eds.), *Directions in Sociolinguistics*, pp. 213–50. New York: Holt.

Escandell Vidal, Victoria (1998) Intonation and procedural encoding: the case of Spanish interrogatives. In V. Rouchota and A. Jucker (eds.), *Current Issues in Relevance Theory*, pp. 169–204. Amsterdam: Benjamins.

Escandell Vidal, Victoria (2004) Norms and principles: putting social and cognitive pragmatics together. In R. Márquez-Reiter and M.E. Placencia (eds.), *Current Trends in the Pragmatics of Spanish*, pp. 347–71. Amsterdam: Benjamins.

Eschenberg, A. (2004) The operator projection, Omaha and diachronic syntax in RRG. In Nolan (2004a: 72–91).

Eschenberg, A. (2005) The article system of UmoNhoN (Siouan). PhD dissertation, City University at Buffalo.

Evans, Nicholas D. (2003) *Bininj Gun-Wok: A Pan-Dialectal Grammar of Mayali, Kunwinjku* and *Kune* (2 vols). Canberra: Australian National University.

Evans, Nicolas D., Duncan Brown, and Greville Corbett (2001) Dalabon pronominal prefixes and the typology of syncretism: a Network Morphology analysis. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2000*, pp. 187–231. Dordrecht: Kluwer.

Evans, Nicholas D., and Toshiki Osada (2005) Mundari: the myth of a language without word classes. *Linguistic Typology* 9: 351–90.

Evans, Nicholas D., and David Wilkins (2000) In the mind's ear: the semantic extensions of perception verbs in Australian languages. *Language* 76: 546–92.

Everett, Dan (2002) Towards an RRG theory of morphology. Lecture delivered at the 2002 International Conference on Role and Reference Grammar (University of La Rioja).

Everett, Dan (2008) Wari 'intentional state constructions'. In Van Valin (2008b: 381–409).

Falk, Yehuda (2001) *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford, CA: CSLI Publications.

Falk, Yehuda (2006) *Subjects and Universal Grammar: An Explanatory Theory*. Cambridge: Cambridge University Press.

Farkas, Donka (2002) Specificity distinctions. *Journal of Semantics* 19: 1–31.

Farkas, Donka, and Henriëtte de Swart (2003) *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. Stanford, CA: CSLI Publications.

Farkas, Donka, and Henriëtte de Swart (2007) Article choice in plural generics. *Lingua* 117: 1657–76.

Farkas, Donka, and Henriëtte de Swart (2009) Stability and variation in article choice: generic and non-generic contexts. In Monika Rathert and Anastasia Giannakidou (eds.), *Quantification, Definiteness and Nominalization*, pp. 188–213. Oxford: Oxford University Press.

Farrell, Patrick (2005) *Grammatical Relations*. Oxford: Oxford University Press.

Fauconnier, Gilles (1985) *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, MA/London: MIT Press/Bradford.

Fauconnier, Gilles, and Mark Turner (2002) *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Feilke, Helmuth (1998) Idiomatische Prägung. In Ingrid Barz and Günther Öhlschläger (eds.), *Zwischen Grammatik und Lexikon*, pp. 69–80. Tübingen: Niemeyer.

Feilke, Helmuth (2004) Kontext—Zeichen—Kompetenz: Wortverbindungen unter sprachtheoretischem Aspekt. In Kathrin Steyer (ed.), *Wortverbindungen—mehr oder weniger fest: Jahrbuch 2003 des Instituts für Deutsche Sprache*, pp. 41–64. Berlin: de Gruyter.

Feldman, Jerome (2006) *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, MA: MIT Press.

Feldman, Jerome, and Srini Narayanan (2004) Embodied meaning in a neural theory of language. *Brain and Language* 89: 385–92.

Ferreira, Fernanda, Kiel Christianson, and Andrew Hollingworth (2001) Misinterpretations of garden-path sentences: implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research* 30: 3–20.

Ferrer i Cancho, Ramon (2004) Euclidean distance between syntactically linked words. *Physical Review E* 70: 056135.

Festschrift, Jacob Wackernagel zur Vollendung des 70. Lebensjahres. Göttingen: Vandenhoeck & Ruprecht.

Fidelholz, James (1975) Word frequency and vowel reduction in English. *Chicago Linguistic Society* 11: 200–213.

Fillmore, Charles J. (1968) The case for case. In Emmon Bach and Robert T. Harms (eds.), *Universals in Linguistic Theory*, pp. 1–88. New York: Holt, Rinehart & Winston.

Fillmore, Charles J. (1977a) Scenes-and-frames semantics. In Antonio Zampoli (ed.), *Linguistic Structures Processing*, pp. 55–81. Amsterdam: North-Holland.

Fillmore, Charles J. (1977b) The case for case reopened. In P. Cole and J. Sadock (eds.), *Grammatical Relations*, pp. 59–81. New York: Academic Press.

Fillmore, Charles J. (1982) Frame semantics. In Linguistics Society of Korea (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co.

Fillmore, Charles J. (1985) Frames and the semantics of understanding. *Quaderni di semantica* 6.2: 222–54.

Fillmore, Charles J. (1986) Pragmatically controlled zero anaphora. *Berkeley Linguistics Society* 12: 95–107.

Fillmore, Charles J. (1988) The mechanisms of Construction Grammar. *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, pp. 35–55.

Fillmore, Charles J. (1999) Inversion and constructional inheritance. In G. Webelhuth, J.-P. Koenig, and A. Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, pp. 113–28. Stanford: CSLI Publications.

Fillmore, Charles J. (2003) Double-decker definitions: the role of frames in meaning explanations. *Sign Language Studies* (Gallaudet University) 3: 263–95.

Fillmore, Charles J., and Beryl T. S. Atkins (1992) Towards a frame-based organization of the lexicon: the semantics of RISK and its neighbors. In A. Lehrer and E. Kitay (eds.), *Frames, Fields, and Contrast: New Essays in Semantics and Lexical Organization*, pp. 75–102. Hillsdale: Erlbaum.

Fillmore, Charles J., and Beryl T. S. Atkins (1994) Starting where the dictionaries stop: the challenge for computational lexicography. In Beryl T. S. Atkins and A. Zampolli (eds.), *Computational Approaches to the Lexicon*. Oxford: Clarendon Press, pp. 349–93.

Fillmore, Charles J., and Beryl T. S. Atkins (1997) FrameNet and lexicographic relevance. *Proceedings of the First International Conference on Language Resources and Evaluation* (Granada).

Fillmore, Charles J., and Paul Kay (1995) Construction Grammar coursebook. MS, University of California, Berkeley.

Fillmore, Charles J., Paul Kay, and Mary C. O'Connor (1988) Regularity and idiomaticity in grammatical constructions: the case of *let alone*. *Language* 64: 501–38.

Fillmore, Charles J., Paul Kay, Ivan Sag, and Laura A. Michaelis (forthcoming) *Construction Grammar*. Stanford, CA: CSLI Publications.

Fine, Jonathan (2006) *Language in Psychiatry: A Handbook of Clinical Practice*. London: Equinox.

Finkbeiner, M., J. Nicol, K. Nakumura, and D. Greth (2002) The role of language in memory for actions. *Journal of Psycholinguistic Research* 31: 447–57.

Finn, Riona, Mary Hearne, Andy Way, and Josef van Genabith (2006) GF-DOP: Grammatical Feature Data-Oriented Parsing. *Proceedings of the 11th International Conference on LFG* (Konstanz).

Firth, John R. (1968) A new approach to grammar. In F. R. Palmer (ed.), *Selected Papers of J. R. Firth, 1952–1959*, pp. 114–25. Bloomington: Indiana University Press.

Fischer, Klaus (1997) *German–English Verb Valency: A Contrastive Analysis*. Tübingen: Narr.

Fischer, Klaus (1999) Verb valency: an attempt at conceptual clarification. *Web Journal of Modern Language Linguistics 4–5*. (http://wjmll.ncl.ac.uk/issue04–05/fischer.htm)

Fischer, Klaus (2003) Verb, Aussage, Valenzdefinition und Valenzrealisierung: auf dem Weg zu einer typologisch adäquaten Valenztheorie. In Klaas Willems, Ann Coene, and Jeroen Van Pottelberge (eds.), *Valenztheorie: Neuere Perspektiven*, pp. 14–64. Ghent: Akademia Press.

Fischer, Klaus (2007) Temporary ambiguity of German and English term complements. In Thomas Herbst and Katrin Götz-Votteler (eds.), *Valency: Theoretical, Descriptive and Cognitive Issues*, pp. 229–52. Berlin: Mouton de Gruyter.

Fischer, Olga, Muriel Norde, and Harry Perridon (eds.) (2004) *Up and Down the Cline: The Nature of Grammaticalization*. Amsterdam: Benjamins.

Fischer, Olga, Anette Rosenbach, and Dieter Stein (eds.) (2000) *Pathways of Change: Grammaticalization in English*. Amsterdam: Benjamins.

Fiser, József, and Richard N. Aslin (2002) Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences* 99: 15822–6.

Fišiak, Jaček (ed.) (1985) *Historical Semantics, Historical Word Formation.* Berlin: Mouton de Gruyter.

Fitch, Tecumseh, Marc Hauser, and Noam Chomsky (2005) The evolution of the language faculty: clarifications and implications. *Cognition* 97: 179–210.

Fleischman, Suzanne (1982) *The Future in Thought and Language: Diachronic Evidence from Romance.* Cambridge: Cambridge University Press.

Flemming, Edward S. (1995) Auditory representations in phonology. Doctoral dissertation, University of California, Los Angeles.

Flickinger, Daniel (1987) Lexical rules in the hierarchical lexicon. PhD dissertation, Stanford University.

Flier, Michael S. (ed.) (1974) *Slavic Forum: Essays in Linguistics and Literature*. The Hague: Mouton.

Fodor, Janet Dean (1978) Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9: 427–73.

Fodor, Janet Dean (1989) Empty categories in sentence processing.*Language and Cognitive Processes* 4: 155–209.

Fodor, Jerry A. (1983) *The Modularity of Mind*. Cambridge, MA: MIT Press.

Foley, William, and R. D. Van Valin, Jr. (1977) On the viability of the notion of 'subject' in universal grammar. *Berkeley Linguistics Society* 3: 293–320.

Foley, William, and R. D. Van Valin, Jr. (1984) *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Fong, Sandiway, and Robert Berwick (2008) Treebank parsing and knowledge of language: a cognitive perspective. *Proceedings CogSci 2008* (Washington, DC).

Fontenelle, Thierry (ed.) (2003) *International Journal of Lexicography* 16.3 (special issue on FrameNet).

Foulkes, Paul, and Gerard Docherty (2006) The social life of phonetics and phonology. *Journal of Phonetics* 34: 409–38.

Fox, Barbara A., and Sandra A. Thompson (1990) A discourse explanation of the grammar of relative clauses in English conversation. *Language* 66: 297–316.

Fox, Danny (2000) *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.

Fox-Keller, Evelyn (2002) *Making Sense of Life*. Cambridge, MA: Harvard University Press.

Frajzyngier, Zygmunt (1996) *Grammaticalization of the Complex Sentence: A Case Study in Chadic*. Amsterdam: Benjamins.

Frajzyngier, Zygmunt, and Erin Shay (2003) *Explaining Language Structure through Systems Interaction*. Amsterdam: Benjamins.

Frampton, John, and Sam Gutmann (2002) Crash-proof syntax. In Samuel D. Epstein and T. Daniel Seely (eds.), *Explanation and Derivation in the Minimalist Program*, pp. 90–105. Oxford: Blackwell.

Francis, G., S. Hunston, and E. Manning (eds.) (1996) *Collins COBUILD Grammar Patterns 1: Verbs.* London: HarperCollins.

Francis, G., S. Hunston, and E. Manning (eds.) (1998) *Collins COBUILD Grammar Patterns 2: Nouns and Adjectives.* London: HarperCollins.

Frank, Anette, Tracy Holloway King, Jonas Kuhn, and John T. Maxwell, III (1998) Optimality Theory style constraint ranking in large-scale LFG grammars. In Butt and King (1998b), pp. 367–97.

Frank, Anette, and Josef van Genabith (2001) LL-based semantics construction for LTAG—and what it teaches us about the relation between LFG and LTAG. In Butt and King (2001a: 104–26).

Frank, Anette, and Annie Zaenen (2002) Tense in LFG: syntax and morphology. In Hans Kamp and Uwe Reyle (eds.), *How Do We Say WHEN It Happens? Contributions to the Theory of Temporal Reference in Natural Language*. Tübingen: Niemeyer.

Franklin, K. (1971) *A Grammar of Kewa, New Guinea*. Canberra: Australian National University.

Frascarelli, Mara, and Hinterhölzl, Roland (2007) Types of topics in German and Italian. In S. Winkler and K. Schwabe (eds.), *On Information Structure, Meaning and Form*, pp. 87–116. Amsterdam: Benjamins.

Frascarelli, Mara, and Annarita Puglielli (forthcoming) Focus in the Force-Fin system: information structure in Cushitic languages. In Enoch Aboh, K. Hartmann, and M. Zimmermann (eds.), *Focus Strategies: Evidence from African Languages*. Berlin: Mouton de Gruyter.

Frawley, William (1992) *Linguistic Semantics*. Hillsdale, NJ: Erlbaum.

Frazier, Lyn (1987) Sentence processing: a tutorial review. In M. Coltheart (ed.), *Attention and Performance XII: The Psychology of Reading*, pp. 559–86. Mahwah, NJ: Erlbaum.

Frazier, Lyn (1998) Getting there (slowly). *Journal of Psycholinguistic Research* 27: 123–46.

Frazier, Lyn, and Charles Clifton (1989) Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4: 93–126.

Frazier, Lyn, and Charles Clifton (1996) *Construal*. Cambridge, MA: MIT Press.

Freidin, Robert, and Jean-Roger Vergnaud (2001) Exquisite connections: some remarks on the evolution of linguistic theory. *Lingua* 111: 639–66.

Fried, Mirjam, and Hans C. Boas (eds.) (2005) *Grammatical Constructions: Back to the Roots*. Amsterdam: Benjamins.

Friedman, Lynn A. (ed.) (1977) *On the Other Hand: New Perspectives on American Sign Language*. New York: Academic Press.

Friginal, Eric (2008) Linguistic variation in the discourse of outsourced call centers. *Discourse Studies* 10: 715–36.

Friginal, Eric (2009) *The Language of Outsourced Call Centers*. Amsterdam: Benjamins.

Frishberg, Nancy (1975) Arbitrariness and iconicity: historical change in American Sign Language. *Language* 51: 676–710.

Frith, Uta, and Christopher D. Frith (2003) Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 358: 459–73.

Fujii, Seiko (2005) An analysis of communication verbs in the Japanese FrameNet (in Japanese) *Proceedings of the Annual Meeting of the Japanese Cognitive Linguistics Association* (*JCLA* 5), pp. 625–8.

Fukui, Naoki (1995) *Theory of Projection in Syntax*. Stanford, CA: CSLI Publications.

Fukui, Naoki (2006) *Theoretical Comparative Syntax*. London: Routledge.

Furuya, Kaori (2008) DP hypothesis for Japanese 'bare' noun phrases. *University of Pennsylvania Working Papers in Linguistics* 14: 149–62.

Gahl, Susanne, and Susan Garnsey (2004) Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. *Language* 80: 748–75.

Gahl, Susanne, and Alan Yu (eds.) (2006) Special issue on exemplar-based models in linguistics. *Linguistic Review* 23.3.

Gaifman, Haim (1965) Dependency systems and phrase structure systems. *Information and Control* 8: 304–37.

Gair, James and John Paolillo (1997) *Sinhala*. Munich: Lincom-Europa.

Galilei, Galileo (1962[1632]) *Dialogue Concerning the Two Chief World Systems*. Berkeley: University of California Press.

Galilei, Galileo (1974[1638]) *Two New Sciences*. Toronto: Wall & Thompson.

Gallese, Vittorio, and George Lakoff (2005) The brain's concepts: the role of the sensory-motor system in reason and language. *Cognitive Neuropsychology* 22: 455–79.

Gansel, Christina (2003) Valenz und Kognition. In Ágel et al. (2003: 422–44).

Gärdenfors, Peter (2000) *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press.

Gardner, Beatrice T., and R. A. Gardner (1971) Two-way communication with an infant chimpanzee. In Allan Schrier and Fred Stollnitz (eds.), *Behavior of Non-human Primates*, vol. 4, pp. 117–84. New York: Academic Press.

Garnsey, Susan, Michael Tanenhaus, and Robert Chapman (1989) Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18: 51–60.

Gathercole, Susan E., and Alan D. Baddeley (1993) *Working Memory and Language*. Hillsdale, NJ: Erlbaum.

Gawron, Jean Mark (1988) *Lexical Representations and the Semantics of Complementation*. New York: Garland.

Gawron, Jean Mark (forthcoming) Frame semantics. In Claudia Maienborn, Klaus von Heusinger, and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.

Gazdar, Gerald (1979) *Pragmatics: Implicature, Presupposition and Logical Form.* London: Academic Press.

Gazdar, Gerald (1981) Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12: 155–84.

Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985) *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

Geary, D. C. (2005) *The Origin of Mind: The Evolution of Brain, Cognition and General Intelligence*. Washington, DC: American Psychological Association.

Geeraerts, Dirk, and Hubert Cuyckens (2007) *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press.

Geisler, Christer (2002) Investigating register variation in nineteenth-century English: a multi-dimensional comparison. In R. Reppen, S. M. Fitzmaurice, and D. Biber (eds.), *Using Corpora to Explore Linguistic Variation*, pp. 249–71. Amsterdam: Benjamins.

Gentner, Deidre and Susan Goldin-Meadow (eds.) (2003) *Language in Mind: Advances in the Study of Language and Thought.* Cambridge, MA: MIT Press.

Gernsbacher, Morton Ann (1990) *Language Comprehension as Structure Building*. Hillsdale, NJ: Erlbaum.

Geschwind, Norman (1970) The organization of language and the brain. *Science* 170: 940–44.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen, and Kevin Russell (2004) Contrastive focus reduplication in English (the 'salad-salad' paper). *Natural Language and Linguistic Theory* 22: 307–57.

Gibson, Edward (2002) The influence of referential processing on sentence complexity *Cognition* 85: 79–112.

Ginzburg, Jonathan, and Ivan A. Sag (2000) *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. Stanford: CSLI Publications.

Girard, Jean-Yves (1987) Linear logic. *Theoretical Computer Science* 50: 1–102.

Girard, Jean-Yves, Paul Taylor, and Yves Lafont (1989) *Proofs and Types.* Cambridge: Cambridge University Press.

Gisborne, Nikolas (2001) The stative/dynamic contrast and argument linking. *Language Sciences* 23: 603–37.

Giusti, Giuliana (2002) The functional structure of noun phrases: a bare phrase structure approach. In G. Cinque (ed.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures*, vol. 1, pp. 54–90. New York: Oxford University Press.

Givón, T. (1971) Historical syntax and synchronic morphology: an archaeologist's field trip. *Chicago Linguistics Society* 7: 394–415.

Givón, T. (1975a) Serial verbs and syntactic change: Niger-Congo. In Li (1975: 47–112).

Givón, T. (1975b) Topic, pronoun and grammatical agreement. In Li (1975: 149–88).

Givón, T. (1979) *On Understanding Grammar*. New York: Academic Press.

Givón, T. (1980) The binding hierarchy and the typology of complements. *Studies in Language* 4: 333–77.

Givón, T. (1981) On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 2: 35–53.

Givón, T. (1984) *Syntax: A Functional-Typological Introduction*, vol. 1. Amsterdam: Benjamins.

Givón, T. (1989) *Mind, Code and Context: Essays in Pragmatics*. Hillsdale, NJ: Erlbaum.

Givón, T. (1990a) *Syntax: A Functional-Typological Introduction*, vol. 2. Amsterdam: Benjamins.

Givón, T. (1990b) Natural language learning and organized language teaching. In H. Burmeister and P. Rounds (eds.), *Variability in Second Language Acquisition: Proceedings of the 10th Second Language Research Forum* (SLRF), pp. 61–84. Eugene: University of Oregon.

Givón, T. (1991) The grammar of referential coherence as mental processing instructions. *Linguistics* 30: 5–55.

Givón, T. (1995) *Functionalism and Grammar*. Amsterdam: Benjamins.

Givón, T. (2001) *Syntax: An Introduction* (2 vols). Amsterdam: Benjamins.

Givón, T. (2002) *Bio-Linguistics*. Amsterdam: Benjamins.

Givón, T. (2005) *Context as Other Minds: The Pragmatics of Cognition and Communication*. Amsterdam: Benjamins.

Givón, T. (2007) The ontogeny of complex verb phrases: how children learn to negotiate fact and desire. TS, University of Oregon.

Givón, T. (2008) *The Genesis of Syntactic Complexity*. Amsterdam: Benjamins.

Givón, T., and Bertram F. Malle (eds.) (2002) *The Evolution of Language out of Pre-language*. Amsterdam: Benjamins.

Givón, T., and Sue Savage-Rumbaugh (2006) Can apes learn grammar? A short detour into language evolution. In Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura, and Seyda Ozcaliskan (eds.), *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin*. London: Psychology Press, pp. 299–309.

Gladkova, Anna (2007a) Russian emotions, attitudes and values: selected topics in cultural semantics. PhD thesis, Australian National University.

Gladkova, Anna (2007b) Universal and language-specific aspects of propositional attitudes : Russian vs. English. In Andrea C. Schalley and Drew Khlentzos (eds.), *Mental States*, vol. 2: *Language and Cognitive Structure*, pp. 61–83. Amsterdam: Benjamins.

Glenberg, Arthur, and Michael Kaschak (2002) Grounding language in action. *Psychonomic Bulletin and Review* 9: 558–65.

Gnanadesikan, Amalia (2004) Markedness and faithfulness constraints in child phonology. In René Kager, Joe Pater, and Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press. Available on Rutgers Optimality Archive.

Goddard, Cliff (1998) *Semantic Analysis: A Practical Introduction*. Oxford: Oxford University Press.

Goddard, Cliff (2001) Lexico-semantic universals: a critical overview. *Linguistic Typology* 5: 1–66.

Goddard, Cliff (2002) Ethnosyntax, ethnopragmatics, sign-functions, and culture. In N. J. Enfield (ed.), *Ethnosyntax*, pp. 52–73. Oxford: Oxford University Press.

Goddard, Cliff (2003a) 'Thinking' across languages and cultures: six dimensions of variation. *Cognitive Linguistics* 14: 109–40.

Goddard, Cliff (2003b) Whorf meets Wierzbicka: variation and universals in language and thinking. *Language Sciences* 25: 393–432.

Goddard, Cliff (2003c) Dynamic *ter-* in Malay (*Bahasa Melayu*): a study in grammatical polysemy. *Studies in Language* 27: 287–322.

Goddard, Cliff (2006a) Ethnopragmatics: a new paradigm. In Goddard (2006b: 1–30).

Goddard, Cliff (ed.) (2006b) *Ethnopragmatics: Understanding Discourse in Cultural Context*. Berlin: Mouton de Gruyter.

Goddard, Cliff (ed.) (2008) *Cross-Linguistic Semantics.* Amsterdam: Benjamins.

Goddard, Cliff, and Anna Wierzbicka (eds.) (1994) *Semantic and Lexical Universals: Theory and Empirical Findings.* Amsterdam: Benjamins.

Goddard, Cliff, and Anna Wierzbicka (1997) Discourse and culture. In Teun A. van Dijk (ed.), *Discourse Studies: A Multidisciplinary Introduction*, vol. 2: *Discourse as Social Interaction*, pp. 231–57. London: Sage.

Goddard, Cliff, and Anna Wierzbicka (2002a) Semantic primes and universal grammar. In Goddard and Wierzbicka (2002b: vol. 1, 41–85).

Goddard, Cliff, and Anna Wierzbicka (eds.) (2002b) *Meaning and Universal Grammar: Theory and Empirical Findings* (2 vols). Amsterdam: Benjamins.

Goddard, Cliff, and Anna Wierzbicka (eds.) (2004) *Cultural Scripts*. Special issue of *Intercultural Pragmatics* 1.2.

Goddard, Cliff, and Anna Wierzbicka (2007) NSM analyses of the semantics of physical qualities: *sweet*, *hot*, *hard*, *heavy*, *rough*, *sharp* in cross-linguistic perspective. *Studies in Language* 31: 765–800.

Goddard, Cliff, and Anna Wierzbicka (2009) Contrastive semantics of physical activity verbs: 'cutting' and 'chopping' in English, Polish, and Japanese. *Language Sciences* 31: 60–96.

Goddard, Cliff, and Anna Wierzbicka (forthcoming) *Men*, *women* and *children*: the conceptual semantics of basic social categories.

Goffman, Erving (1974) *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper & Row.

Goldberg, Adele E. (1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. (2002) Surface generalizations: an alternative to alternations. *Cognitive Linguistics* 13: 327–56.

Goldberg, Adele E. (2005) Constructions, lexical semantics and the Correspondence Principle: accounting for generalizations and subregularities in the realization of arguments. In N. Erteschik-Shir and T. Rapoport (eds.), *The Syntax of Aspect*, pp. 211–236. Oxford: Oxford University Press.

Goldberg, Adele E. (2006) *Constructions at Work: The Nature of Generalizations in Language.* Oxford: Oxford University Press.

Goldberg, Adele E., Devin Casenhiser, and Nitya Sethuraman (2004) Learning argument structure generalizations. *Cognitive Linguistics* 14: 289–316.

Goldberg, Adele E., and Alex del Giudice (2005) Subject–auxiliary inversion: a natural category. *Linguistic Review* 22: 411–28.

Goldinger, Stephen D. (1996) Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22: 1166–83.

Goldinger, Stephen D. (2000) The role of perceptual episodes in lexical processing. In Anne Cutler, James McQueen, and Rian Zondervan (eds.), *Proceedings of Spoken Word Access Processes*, pp. 155–8. Nijmegen: Max Planck Institute for Psycholinguistics.

Good, Jeff (1983) Iconic and economic motivation. *Language* 59: 781–819.

Good, Jeff (ed.) (2008) *Linguistic Universals and Language Change.* Oxford: Oxford University Press.

Goodale, Melvin A. (2000) Perception and action in the human visual system. In M. S. Gazzaniga (ed.), *The New Cognitive Neuroscience*, 2nd edn, pp. 365–77. Cambridge, MA: MIT Press.

Goodman, Joshua (2003) Efficient parsing of DOP with PCFG-reductions. In R. Bod, R. Scha, and K. Sima'an (eds.), *Data-Oriented Parsing*, pp. 125–46. Stanford, CA: CSLI Publications.

Goodman, Judith, Philip Dale, and Ping Li (2008) Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language* 35: 515–31.

Goodwin, Brian C., and Lynn E. Trainor (1983) The ontogeny and phylogeny of the Pentadactyl limb. In Brian C. Goodwin, Nigel Holder, and Christopher C. Wylie (eds.), *Development and Evolution*, pp. Cambridge: Cambridge University Press, pp. 75–98.

Goossens, Louis (1990) Metaphtonymy: the interaction of metaphor and metonymy in expressions for linguistic action. *Cognitive Linguistics* 1: 323–40.

Gopnik, Alison, and Henry M. Wellman (1992) Why the child's theory of mind is really a theory. *Mind and Language* 7: 145–71.

Gordon, Matthew (1999) Syllable weight: phonetics, phonology, and typology. Doctoral dissertation, University of California, Los Angeles.

Gould, Stephen Jay (1983) *Hen's Teeth and Horse's Toes*. New York: Norton.

Gould, Stephen Jay (1989) *Wonderful Life*. New York: Norton.

Gould, Stephen Jay (2002) *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.

Gouskova, Maria (2003) Deriving economy: syncope in Optimality Theory. Doctoral dissertation, University of Massachusetts, Amherst. Available on the Rutgers Optimality Archive, ROA 610, http://roa.rutgers.edu

Gouskova, Maria (2004) Relational hierarchies in Optimality Theory: the case of syllable contact. *Phonology* 21: 201–50.

Graybiel, Ann M. (1998) The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory* 70: 119–36.

Green, Georgia, M. (1996) *Pragmatics and Natural Language Understanding*, 2nd edn. Mahwah, NJ: Erlbaum.

Greenberg, Joseph H. (1963) *Universals of Language*. Cambridge, MA: MIT Press.

Greenberg, Joseph H. (1966a) Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg (1966b: 73–113).

Greenberg, Joseph H. (1966b) *Language Universals: With Special Reference to Feature Hierarchies*. The Hague: Mouton.

Greenberg, Joseph H. (1969) Some methods of dynamic comparison in linguistics. In Jan Puhvel (ed.), *Substance and Structure of Language*, pp. 147–203. Berkeley: University of California Press.

Greenberg, Joseph H. (1978a) Diachrony, synchrony and language universals. In Joseph Greenberg, Charles Ferguson, and Edith Moravcsik (eds.), *Universals of Human Language*, vol. 1, pp. 61–91. Stanford, CA: Stanford University Press.

Greenberg, Joseph H. (1978b) How does a language acquire gender markers? In Joseph Greenberg, Charles Ferguson, and Edith Moravcsik (eds.), *Universals of Human Language*, vol. 3, pp. 47–82. Stanford, CA: Stanford University Press.

Greenfield, Patricia M. (1991) Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavior and Brain Science* 14: 531–95.

Gregory, Michelle, William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky (1999) The effects of collocational strength and contextual predictibility in lexical production. *Chicago Linguistic Society* 35: 151–66.

Greiffenhagen, Christian, and Wes Sharrock (2007) Linguistic relativism: logic, grammar, and arithmetic in cultural comparison. *Language and Communication* 27: 81–107.

Grenander, Ulf (1967) *Syntax-Controlled Probabilities*. Technical Report, Division of Applied Mathematics, Brown University, Providence, RI.

Grewendorf, Günther (2002) Left dislocation as movement. In S. Manck and J. Mittelstaedt (eds.), *Georgetown University Working Papers in Theoretical Linguistics*, pp. 31–81.

Grice, H. Paul (1975) Logic and conversation. In P. Cole and J. L. Morgan (eds.), *Speech Acts*, pp. 41–58. New York: Academic Press.

Grice, H. Paul (1978) Further notes on logic and conversation. In P. Cole (ed.), *Pragmatics*. New York: Academic Press. Reprinted in Grice (1989: 41–57).

Grice, H. Paul (1989) *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Grice, H. Paul (2001) *Aspects of Reason*, ed. R. Warner. Oxford: Clarendon Press.

Gries, Stefan (2005) Syntactic priming: a corpus-based approach. *Journal of Psycholinguistic Research* 34: 365–99.

Grillo, Nino (2008) Generalized minimality: syntactic underspecification in Broca's Aphasia. PhD dissertation, University of Utrecht/University of Siena.

Grimes, Barbara (1996) *Ethnologue*. Dallas, TX: Summer Institute of Linguistics.

Grinder, John (1970) Super equi-NP deletion. *Proceedings of the Sixth Meeting of the Chicago Linguistic Society*, pp. 297–317.

Groefsema, Marjolein (1995) *Can*, *may*, *must* and *should*: a relevance-theoretic account. *Journal of Linguistics* 31: 53–79.

Gruber, Jeffrey (1965) Studies in lexical relations. PhD dissertation, MIT. Published in J. Gruber, *Lexical Structures in Syntax and Semantics*, Amsterdam: North-Holland, 1976.

Guasti, Maria Teresa (2002) *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press.

Guerrero, Lilián, Sergio Ibáñez, and Valeria Belloro (eds.) (2009) *Studies in Role and Reference Grammar*. Mexico: Universidad Nacional Autónoma de México.

Guerrero Valenzuela, L., and R. D. Van Valin, Jr. (2004) Yaqui and the analysis of primary object languages. *International Journal of Applied Linguistics* 70: 290–319.

Guest, E. (2008) Parsing for Role and Reference Grammar. In Van Valin (2008b: 435–53).

Guest, E., and S. Brown (2007) Using Role and Reference Grammar to support computer-assisted assessment of free text answers. MS, Leeds Metropolitan University.

Guillaume, P. (1973[1927]) The development of formal elements in the child's speech. In Charles A. Ferguson and Daniel I. Slobin (eds.), *Studies of Child Language Development*, pp. 240–51. New York: Holt, Rinehart & Winston.

Gumperz, John Joseph, and Stephen C. Levinson (1996) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Gutt, Ernst-August (2000) *Translation and Relevance: Cognition and Context*, 2nd edn. Manchester: St Jerome Publishing.

Habermann, Mechthild (2007) Aspects of a diachronic valency syntax of German. In Thomas Herbst and Katrin Götz-Votteler (eds.), *Valency: Theoretical, Descriptive and Cognitive Issues*, pp. 85–100. New York: Mouton de Gruyter.

Haegeman, Liliane (1989) *Be going to* and *will*: a pragmatic account. *Journal of Linguistics* 25: 291–319.
Haegeman, Liliane (1994) *Introduction to Government and Binding Theory*, 2nd edn. Oxford: Blackwell.
Haegeman, Liliane (2006) Argument fronting in English, Romance ClLD and left periphery. In R. Zanuttini, H. Campos, E. Herburger, and P. Portner (eds.), *Crosslinguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture*, pp. 27–52. Washington, DC: Georgetown University Press.
Hagège, Claude (1985) *L'Homme de paroles: contribution linguistique aux sciences humaines*. Paris: Fayard.
Hagège, Claude (1993) *The Language Builder: An Essay on the Human Signature in Linguistic Morphogenesis*. Amsterdam: Benjamins.
Haiman, John (1985) *Natural Syntax*. Cambridge: Cambridge University Press.
Haiman, John (1994) Ritualization and the development of language. In William Pagliuca (ed.), *Perspectives on Grammaticalization*, pp. 3–28. Philadelphia: Benjamins.
Haiman, John (1997) Repetition and identity. *Lingua* 100: 57–70.
Hale, Kenneth, and Samuel J. Keyser (1993) On argument structure and the lexical expression of syntactic relations. In Kenneth Hale and Samuel J. Keyser (eds.), *The View from Building 20*, pp. 53–110. Cambridge, MA: MIT Press.
Hale, Kenneth, and Samuel J. Keyser (2002) *Prolegomena to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
Hale, Mark, and Charles Reiss (2000) 'Substance abuse' and 'dysfunctionalism': current trends in phonology. *Linguistic Inquiry* 31: 157–69.
Halle, Morris (1962) Phonology in generative grammar. *Word* 18: 54–72.
Halle, Morris, and Alec Marantz (1993) Distributed morphology and the pieces of inflection. In Kenneth Hale and Samuel Keyser (eds.), *The View From Building 20*, pp. 111–76. Cambridge, MA: MIT Press.
Halle, Morris, and Jean-Roger Vergnaud (1987) *An Essay on Stress*. Cambridge, MA: MIT Press.
Halliday, M. A. K. (1961) Categories of the theory of grammar. *Word* 17: 241–92.
Halliday, M. A. K. (1966) Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, and R. H. Robins (eds.) *In Memory of J. R. Firth*, pp. 148–62. London: Longmans.
Halliday, M. A. K. (1971) Linguistic function and literary style: an inquiry into the language of William Golding's *The Inheritors*. In Seymour Chatman (ed.), *Literary Style: A Symposium*, pp. 330–68. Oxford: Oxford University Press.
Halliday, M. A. K. (1977) Ideas about language. In *Aims and Perspectives in Linguistics*. Applied Linguistics Association of Australia: Occasional Papers Number 1.
Halliday, M. A. K. (1978) *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
Halliday, M. A. K. (1982) The de-automatization of grammar: from Priestley's *An Inspector Calls*. In John M. Anderson (ed.), *Language Form and Linguistic Variation: Papers dedicated to Angus McIntosh*, pp. 129–59. Amsterdam: Benjamins.
Halliday, M. A. K. (1994) *An Introduction to Functional Grammar*, 2nd edn. London: Arnold.
Halliday, M. A. K. (1996) On grammar and grammatics. In Ruqaiya Hasan, Carmel Cloran. and David Butt (eds.), *Functional Descriptions: Theory in Practice*. Amsterdam: Benjamins, pp. 1–38.

Halliday, M. A. K. (2002) *Linguistics Studies of Text and Discourse*, ed. Jonathan J. Webster. London: Continuum.

Halliday, M. A. K. (2005) Computing meaning: some reflections on past experience and present prospects (1995). In Jonathan J. Webster (ed.), *Computational and Quantitative Studies*, pp. 239–67. London: Continuum.

Halliday, M. A. K., and Ruqaiya Hasan (1976) *Cohesion in English*. London: Longman.

Halliday, M. A. K., and Ruqaiya Hasan (1985) *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Geelong, Vic.: Deakin University Press.

Halliday, M. A. K., and Christian M. I. M. Matthiessen (1999) *Construing Experience through Meaning: A Language-Based Approach to Cognition*. London: Continuum.

Halliday, M. A. K., and Christian M. I. M. Matthiessen (2004) *An Introduction to Functional Grammar*, 3rd edn. London: Arnold.

Halvorsen, Per-Kristian, and Ronald M. Kaplan (1988) Projections and semantic description in Lexical-Functional Grammar. *Proceedings of the International Conference on Fifth Generation Computer Systems*, pp. 1116–22 (Institute for New Generation Systems, Tokyo). Reprinted in Dalrymple et al. (1995: 279–92).

Hamm, Fritz, Hans Kamp, and Michiel van Lambalgen (2006) There is no opposition between Formal and Cognitive Semantics. *Theoretical Linguistics* 32: 1–40.

Hammond, Michael (1999) *The Phonology of English*. Oxford: Oxford University Press.

Hankamer, Jorge, and Judith Aissen (1974) The sonority hierarchy. In Anthony Bruck, Robert Fox, and Michael La Galy (eds.), *CLS 10: Parasession on Natural Phonology*, pp. 131–45. Chicago: Chicago Linguistic Society.

Hanks, William F. (1990) *Referential Practice: Language and Lived Space among the Maya*. Chicago: University of Chicago Press.

Hansell, M. (1993) Serial verbs and complement constructions in Mandarin: a clause linkage analysis. In Van Valin (1993a: 197–233).

Hardin, C. L., and Luisa Maffi (1997) *Color Categories in Thought and Language*. Cambridge: Cambridge University Press.

Harkins, Jean and Anna Wierzbicka (eds.) (2001) *Emotions in Crosslinguistic Perspective*. Berlin: Mouton de Gruyter.

Harley, Trevor (1995) *The Psychology of Language*. Hove: Psychology Press.

Harnish, Robert M. (1976) Logical form and implicature. In T. Beaver, J. Katz, and D. T. Langendoen (eds.), *An Integrated Theory of Linguistic Ability*, pp. 313–92. New York: Crowell.

Harrington, Jonathan (2006) An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics* 34: 439–57.

Harrington, Jonathan, Sallyanne Palethorpe, and Catherine I. Watson (2000) Does the Queen speak the Queen's English? *Nature* 408: 927–8.

Harris, Alice C 2006. In other words: external modifiers in Georgian. *Morphology* 16: 205–29.

Harris, Alice C., and Lyle Campbell (1995) *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Harris, James (1983) *Syllable Structure And Stress in Spanish: A Nonlinear Analysis*. Cambridge, MA: MIT Press.

Harris, Zellig (1954) Distributional structure. *Word* 10: 146–62.

HASADA, RIE (2008) Two 'virtuous emotions' in Japanese: *nasake/joo* and *jihi*. In Cliff Goddard (ed.), *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 331–47. Berlin: Mouton de Gruyter.

HASAN, RUQAIYA (1989) Semantic variation and sociolinguistics. *Australian Journal of Linguistics* 9: 221–75.

HASAN, RUQAIYA, CHRISTIAN MATTHIESSEN, and JONATHAN J. WEBSTER (eds.) (2005) *Continuing Discourse on Language: A Functional Perspective*, vol. 1. London: Equinox.

HASAN, RUQAIYA, CHRISTIAN MATTHIESSEN, and JONATHAN J. WEBSTER (eds.) (2007) *Continuing Discourse on Language: A Functional Perspective*, vol. 2. London: Equinox.

HASPELMATH, MARTIN (1998) Does grammaticalization need reanalysis? *Studies in Language* 22: 315–51.

HASPELMATH, MARTIN (1999a) Why is grammaticalization irreversible? *Linguistics* 37: 1043–68.

HASPELMATH, MARTIN (1999b) Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18: 180–205.

HASPELMATH, MARTIN (2004a) Does linguistic explanation presuppose linguistic description? *Studies in Language* 28: 554–79. (Special issue on linguistic evidence, guest edited by Martina Penke and Anette Rosenbach.)

HASPELMATH, MARTIN (2004b) On directionality in language change with particular reference to grammaticalization. In Fischer et al. (2004: 17–44).

HASPELMATH, MARTIN (2006) Against markedness (and what to replace it with). *Journal of Linguistics* 42: 25–70.

HASPELMATH, MARTIN (2007) Pre-established categories don't exist: consequences for language description and typology. *Linguistic Typology* 111: 119–32.

HASPELMATH, MARTIN (2008a) Creating economical morphosyntactic patterns in language change. In Good (2008: 185–214).

HASPELMATH, MARTIN (2008b) Parametric versus functional explanations of syntactic universals. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, pp. 75–107. Amsterdam: Benjamins.

HASPELMATH, MARTIN (2008c) Ditransitive constructions: towards a new Role and Reference Grammar account? In Van Valin (2008b: 75–100).

HASPELMATH, MARTIN (2008d) Creating economical morphosyntactic patterns in language change. In Good (2008: 185–214).

HASPELMATH, MARTIN (2008e) Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19: 1–33.

HASPELMATH, MARTIN (2009a) Descriptive categories and comparative concepts in cross-linguistic studies. MS, Max Planck Institute for Evolutionary Anthropology.

HASPELMATH, MARTIN (2009b) Descriptive scales versus comparative scales. In Marc Richards, Andrej L. Malchukov, and Ina Bornkessel (eds.), *Scales in Grammatical Theory*.

HASPELMATH, MARTIN (2009c) An empirical test of the agglutination hypothesis. In Elisabetta Magni, Sergio Scalise, and Antonietta Bisetto (eds.), *Universals of Language Today*, pp. 13–29. Dordrecht: Springer.

HASPELMATH, MARTIN, MATTHEW DRYER, DAVID GIL, and BERNARD COMRIE (eds.) (2005) *The World Atlas of Language Structures*. Oxford: Oxford University Press.

HAUK, OLAF, INGRID JOHNSRUDE, and FRIEDEMANN PULVERMÜLLER (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41: 301–7.

HAUSER, MARC D., NOAM CHOMSKY, and W. TECUMSEH FITCH (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* 298: 1569–79.

HAWKINS, JOHN A. (1986) *A Comparative Typology of English and German: Unifying the Contrasts.* London: Croom Helm.

HAWKINS, JOHN A. (1991) On (in)definite articles: implicatures and (un)grammaticality predictions. *Journal of Linguistics* 27: 405–42.

HAWKINS, JOHN A. (1994) *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

HAWKINS, JOHN A. (2001) Why are categories adjacent? *Journal of Linguistics* 37: 1–34.

HAWKINS, JOHN A. (2002) Symmetries and asymmetries: their grammar, typology and parsing. *Theoretical Linguistics* 28: 95–149.

HAWKINS, JOHN A. (2004a) *Efficiency and Complexity in Grammars.* Oxford: Oxford University Press.

HAWKINS, JOHN A. (2004b) *A Performance Theory of Word Order and Constituency*. Cambridge: Cambridge University Press.

HAY, JENNIFER (2001) Lexical frequency in morphology: is everything relative? *Linguistics* 39: 1041–70.

HAY, JENNIFER (2002) From speech perception to morphology: affix ordering revisited. *Language* 78: 527–55.

HAY, JENNIFER, and R. HARALD BAAYEN (2002) Parsing and productivity. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 2001*, pp. 203–36. Dordrecht: Springer.

HAY, JENNIFER, and JOAN BRESNAN (2006) Spoken syntax: the phonetics of giving a hand in New Zealand English.*Linguistic Review* 23: 321–49.

HAY, JENNIFER, and INGO PLAG (2004) What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language and Linguistic Theory* 22, 3: 565–96.

HAYES, BRUCE (1989) The prosodic hierarchy in meter. In Paul Kiparsky and Gilbert Youmans (eds.), *Rhythm and Meter*, pp. 201–60. Orlando, FL: Academic Press.

HAYES, BRUCE (1995) *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

HAYES, BRUCE (1999) Phonetically driven phonology: the role of Optimality Theory and inductive grounding. In Darnell et al. (1999: 243–85).

HAYES, BRUCE (2004) Phonological acquisition in Optimality Theory: the early stages. In René Kager, Joe Pater, and Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, Cambridge: Cambridge University Press.

HAYES, BRUCE, ROBERT KIRCHNER, and DONCA STERIADE (eds.) (2004) *Phonetically-Based Phonology*. Cambridge: Cambridge University Press.

HEARNE, MARY, and KHALIL SIMA'AN (2003) Structured parameter estimation for LFG-DOP using backoff. *Proceedings of Recent Advances in Natural Language Processing.*

HEARNE, MARY, and ANDY WAY (2003) Seeing the wood for the trees: data-oriented translation. *Proceedings of MT Summit IX* (New Orleans).

HEARNE, MARY, and ANDY WAY (2004) Data-oriented parsing and the Penn Chinese Treebank. *Proceedings of the 1st International Joint Conference on Natural Language Processing* (Hainan Island), pp. 406–13.

HEIM, I. (1982) The semantics of definite and indefinite NPs. PhD dissertation, University of Massachusetts at Amherst.

Heine, Bernd (1992) Grammaticalization chains. *Studies in Language* 16: 335–68.
Heine, Bernd (1993) *Auxiliaries: Cognitive Forces and Grammaticalization.* Oxford: Oxford University Press.
Heine, Bernd (1997) *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
Heine, Bernd (2003) Grammaticalization. In Joseph and Janda (2003: 575–601).
Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer (1991) *Grammaticalization: A Conceptual Framework.* Chicago: University of Chicago Press.
Heine, Bernd, and Christa König (2005) Grammatical hybrids: between serialization, compounding and derivation in !Xun. In Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer (eds.), *Morphology and its Demarcations*, pp. 81–96. Amsterdam: Benjamins.
Heine, Bernd, and Tania Kuteva (2002) *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
Heine, Bernd, and Tania Kuteva (2003) Contact-induced grammaticalization. *Studies in Language* 27: 529–72.
Heine, Bernd, and Tania Kuteva (2005) *Language Contact and Grammatical Change.* Cambridge: Cambridge University Press.
Heine, Bernd, and Tania Kuteva (2006) *The Changing Languages of Europe*. Oxford: Oxford University Press.
Heine, Bernd, and Tania Kuteva (2007) *The Genesis of Grammar: A Reconstruction.* Oxford: Oxford University Press.
Heine, Bernd, and Hiroyuki Miyashita (2008) Accounting for a functional category: German *drohen* 'to threaten'. *Language Sciences* 30: 53–101.
Heine, Bernd, and Hiroyuki Miyashita (forthcoming) The intersection between reflexives and reciprocals: a grammaticalization perspective. In König and Gast (forthcoming).
Heine, Bernd, and Mechthild Reh (1982) *Patterns of Grammaticalization in African Languages.* Cologne: Universität zu Köln, Institut für Sprachwissenschaft.
Heine, Bernd, and Mechthild Reh (1984) *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
Heine, Bernd, and Mechthild Reh (2004) *Grammaticalization and Reanalysis in African Languages*, 2nd edn. Hamburg: Buske.
Helbig, Gerhard (1992) *Probleme der Valenz- und Kasustheorie*. Tübingen: Niemeyer.
Helbig, Gerhard, and Wolfgang Schenkel (1969) *Wörterbuch zur Valenz und Distribution deutscher Verben*. 2nd rev. edn 1973, 6th edn 1982, Leipzig: Bibliographisches Institut; 8th edn 1991, Tübingen: Niemeyer.
Helt, Marie E. (2001) A multi-dimensional comparison of British and American spoken English. In Conrad and Biber (2001: 157–70).
Henderson, Mara (2006) Between lexical and lexico-grammatical classification: nominal classification in Sinhala. In Robert Englebretson and Carol Genetti (eds.), *Santa Barbara Papers in Linguistics* 17, pp. 29–48. Santa Barbara: University of California.
Hendriks, Petra, and Helen de Hoop (2001) Optimality Theoretic semantics. *Linguistics and Philosophy* 24: 1–32.
Hengeveld, Kees, and J. Lachlan Mackenzie (2008) *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure.* Oxford: Oxford University Press.
Hengeveld, Kees, and J. Lachlan Mackenzie (2009) Functional Discourse Grammar. In Bernd Heine and Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis.* Oxford: Oxford University Press, pp. 367–400.

Hengeveld, Kees, and Gerry Wanders (2007) Adverbial conjunctions in Functional Discourse Grammar. In Mike Hannay and Gerard J. Steen (eds.), *Structural-Functional Studies in English Grammar: In Honour of Lachlan Mackenzie*, pp. 209–26. Amsterdam: Benjamins.

Herbst, Thomas, David Heath, Ian F. Roe, and Dieter Götz (2004) *A Valency Dictionary of English*. Berlin: Mouton de Gruyter.

Heringer, Hans Jürgen (1984) Neues von der Verbszene. In Gerhard Stickel (ed.), *Pragmatik in der Grammatik: Jahrbuch 1983 des Instituts für deutsche Sprache*, pp. 34–64. Düsseldorf: Schwann.

Heringer, Hans Jürgen (1985) The verb and its semantic power: association as a basis for Valency Theory. *Journal of Semantics* 4: 79–99.

Heringer, Hans Jürgen (1996) *Deutsche Syntax dependentiell*. Tübingen: Stauffenburg.

Heringer, Hans Jürgen (2006) Prinzipien des Valenzwandels. In Ágel et al. (2006: 1447–61).

Heusinger, K. von (1999) Intonation and information structure. Habilitationschrift, University of Konstanz.

Heyes, Cecilia M. (1998) Theory of mind in non-human primates. *Behavioral and Brain Sciences* 21: 101–14.

Higginbotham, James (1985) On semantics. *Linguistic Inquiry* 16: 547–93.

Higgins, D. D. (1923) *How to Talk to the Deaf*. St Louis, MO: Catholic Church at 1118 N. Grand Blvd.

Hill, Jane H., and Bruce Mannheim (1992) Language and world view. *Annual Review of Anthropology* 21: 381–406.

Himmelmann, Nikolaus (1997) *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.

Himmelmann, Nikolaus (2008) Lexical categories and voice in Tagalog. In P. Austin and S. Musgrave (eds.), *Voice and Grammatical Functions in Austronesian Languages*, pp. 247–93. Stanford: CSLI Publications.

Hinds, John (1986) *Japanese*. London: Croom Helm.

Hinzen, Wolfram (2007) *An Essay on Names and Truth*. Oxford: Oxford University Press.

Hirschberg, Julia (1991) *A Theory of Scalar Implicature*. New York: Garland.

Hock, Hans Henrich (1987) Emergent grammar. *Berkeley Linguistics Society* 13: 139–57.

Hock, Hans Henrich (1991) *Principles of Historical Linguistics*, 2nd edn. Berlin: Mouton de Gruyter.

Hock, Hans Henrich (1994) Phonogenesis. In Pagliuca (1994: 29–45).

Hock, Hans Henrich (2001) Grammatical constructions and their discourse origins: prototype or family resemblance– In Pütz et al. (2001: 109–29).

Hockenmaier, Julia, and Mark Steedman (2002) Generative models for statistical parsing with Combinatory Categorial Grammar. *Proceedings of 40th Annual Meeting of the Association for Computational Linguistics*.

Hoffmeister, Philip, T. Florian Jaeger, Ivan Sag, Inbal Arnon, and Neal Snider (2007) Locality and accessibility in *wh*-questions. In S. Featherston and W. Sternefeld (eds.), *Roots: Linguistics in Search of its Evidential Base*, pp. 185–206. Berlin: Mouton de Gruyter.

Hofmann, Thomas (1993) *Realms of Meaning: An Introduction to Semantics*. London: Longman.

Hoijer, Harry (1933) Tonkawa: an Indian language of Texas. In Franz Boas and Harry Hoijer (eds.), *Handbook of American Indian Languages*, vol. 3. Gluckstadt: Augustin.

Hoijer, Harry (1946) Tonkawa. In Harry Hoijer, L. Bloomfield, Mary Haas, A. Halpern, F. K. Li, Stanley Newman, Morris Swadesh, G. L. Trager, C. Voegelin, and Benjamin L. Whorf (eds.), *Linguistic Structures of Native America*. New York: Viking Fund Publications in Anthropology 6, pp. 289–311.

Hoijer, Harry (1949) *An Analytical Dictionary of the Tonkawa Language*. Berkeley: University of California Press.

Holisky, D. (1987) The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71: 103–32.

Holland, John H. (1998) *Emergence: From Chaos to Order*. Reading, MA: Addison-Wesley.

Holton, Gerald (1973) *Thematic Origins of Scientific Thought: Kepler to Einstein*. Cambridge, MA: Harvard University Press.

Hoogweg, Lars (2003) Extending DOP with insertion. In R. Bod, R. Scha, and K. Sima'an (eds.), *Data-Oriented Parsing*, pp. 317–35. Stanford, CA: CSLI Publications.

Hoop, Helen de, and Henriëtte de Swart (2000) Adjunct clauses in optimality theory. *Rivista di linguistica/Italian Journal of Linguistics* 12: 107–27.

Hooper, Joan B. (1976) Word frequency in lexical diffusion and the source of morphophonological change. In William Christie (ed.), *Current Progress in Historical Linguistics*, pp. 96–105. Amsterdam: North-Holland.

Hopper, Paul J. (1987) Emergent grammar. In Jon Aske, Natasha Beery, Laura Michaelis, and Hana Filip (eds.), *Proceedings of the 13th Annual Meeting of the Berkeley Linguistic Society*, pp. 139–57. Berkeley: University of California at Berkeley.

Hopper, Paul J., and Sandra A. Thompson (1980) Transitivity in grammar and discourse. *Language* 56: 251–99.

Hopper, Paul J., and Sandra A. Thompson (1985) The iconicity of the universal categories 'noun' and 'verb'. In John Haiman (ed.), *Iconicity in Syntax*, pp. 151–83. Amsterdam: Benjamins.

Hopper, Paul J. (1991) On some principles of grammaticization. In Traugott and Heine (1991a: 17–35).

Hopper, Paul, and Elizabeth C. Traugott (1993) *Grammaticalization*. Cambridge: Cambridge University Press.

Hopper, Paul J., and Elizabeth C. Traugott (2003) *Grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.

Horn, Laurence R. (1972) On the semantic properties of logical operators in English. PhD dissertation, University of California at Los Angeles.

Horn, Laurence R. (1984) Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (ed.) *Meaning, Form, and Use in Context: Linguistic Applications*, pp. 11–42. Washington, DC: Georgetown University Press.

Horn, Laurence R. (1988) Pragmatic theory. In F. J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, vol. 1, pp. 113–45. Cambridge: Cambridge University Press.

Horn, Laurence R. (1989) *A Natural History of Negation*. Chicago: University of Chicago Press.

Horn, Lawrence R. (2004) Implicature. In Horn and Ward (2004: 3–28).

Horn, Laurence R. (2006) The border wars: a neo-Gricean perspective. In von Heusinger and Turner (2006: 21–48).

Horn, Laurence R. (2007) Neo-Gricean pragmatics: a Manichaean manifesto. In Burton-Roberts (2007: 158–83).

Horn, Lawrence R., and Gregory Ward (eds.) (2004) *The Handbook of Pragmatics*. Oxford: Blackwell.

Hornstein, Norbert (1999) Control and movement. *Linguistic Inquiry* 30: 69–96.

Hornstein, Norbert (2001) *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

Hornstein, Norbert (2005) Introduction to the 2nd edition of Chomsky, *Rules and Representations*. New York: Columbia University Press.

Hornstein, Norbert (2009) *A Theory of Syntax*. Cambridge: Cambridge University Press.

Hornstein, Norbert, Jairo Nunes, and Kleanthes K. Grohmann (2006) *Understanding Minimalism*. Cambridge: Cambridge University Press.

Horwich, Paul (2005) *Reflections on Meaning*. Oxford: Clarendon Press.

Horwood, Graham and Se-Kyung Kim (eds.) 2001. Nasal-only syllables in Tswana. *RuLing Papers 2: Working Papers of Rutgers University*. New Brunswick: Department of Linguistics, Rutgers University. pp. 1–20.

Huang, Yan (1991) A neo-Gricean pragmatic theory of anaphora. *Journal of Linguistics* 27: 301–35.

Huang, Yan (1992) Against Chomsky's typology of empty categories. *Journal of Pragmatics* 17: 1–29.

Huang, Yan (1994) *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. Cambridge: Cambridge University Press.

Huang, Yan (1995) On null subjects and null objects in generative grammar. *Linguistics* 33: 1081–1123.

Huang, Yan (1996) A note on the head-movement analysis of long-distance reflexives. *Linguistics* 34: 833–40.

Huang, Yan (1998) Lecture notes on Lexical Pragmatics. MS, University of Reading and University of Oxford.

Huang, Yan (2000a) *Anaphora: A Cross-linguistic Study*. Oxford: Oxford University Press.

Huang, Yan (2000b) Discourse anaphora: four theoretical models. *Journal of Pragmatics* 32: 151–76.

Huang, Yan (2001) Reflections on theoretical pragmatics. *Waiguoyu* 131: 2–14.

Huang, Yan (2002) Logophoric marking in East Asian languages. In T. Güldemann and M. von Roncador (eds.), *Reported Discourse*, pp. 213–24. Amsterdam: Benjamins.

Huang, Yan (2003) On neo-Gricean pragmatics. *International Journal of Pragmatics* 13: 87–110.

Huang, Yan (2004a) Anaphora and the pragmatics-syntax interface. In Horn and Ward (2004: 288–314).

Huang, Yan (2004b) Neo-Gricean pragmatic theory: looking back on the past; looking ahead to the future. *Waiguoyu* 149: 2–25.

Huang, Yan (2005) Lexical narrowing in English. In Y. J. Chen and Y. N. Leung (eds.), *Selected Papers from the 14th International Symposium on English Teaching*, pp. 55–65. Taipei: Crane Publishing Co.

Huang, Yan (2006a) Neo-Gricean pragmatics. In K. Brown (ed.), *The Encyclopedia of Languages and Linguistics*, 2nd edn, vol. 8, pp. 586–90. New York: Elsevier.

Huang, Yan (2006b) Anaphora, cataphora, exophora, logophoricity. In K. Brown (ed.), *The Encyclopedia of Languages and Linguistics*, 2nd edn, vol. 1, pp. 231–8. New York: Elsevier.

Huang, Yan (2007) *Pragmatics*. Oxford: Oxford University Press.

Huddleston, Rodney, and Geoffrey K. Pullum (2005) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hudson, Richard (1971) *English Complex Sentences: An Introduction to Systemic Grammar.* Amsterdam: North-Holland.

Hudson, Richard (1973) An 'item-and-paradigm' approach to Beja syntax and morphology. *Foundations of Language* 9: 504–48.

Hudson, Richard (1976) *Arguments for a Non-transformational Grammar.* Chicago: Chicago University Press.

Hudson, Richard (1980) *Sociolinguistics*. Cambridge: Cambridge University Press.

Hudson, Richard (1984) *Word Grammar*. Oxford: Blackwell.

Hudson, Richard (1990) *English Word Grammar.* Oxford: Blackwell.

Hudson, Richard (1996) *Sociolinguistics*, 2nd edn. Cambridge: Cambridge University Press.

Hudson, Richard (1997a) Inherent variability and linguistic theory. *Cognitive Linguistics* 8: 73–108.

Hudson, Richard (1997b) The rise of auxiliary *DO*: Verb-non-raising or category-strengthening? *Transactions of the Philological Society* 95: 41–72.

Hudson, Richard (1999) Subject-verb agreement in English. *English Language and Linguistics* 3: 173–207.

Hudson, Richard (2001) Clitics in Word Grammar. *UCL Working Papers in Linguistics* 13: 243–94.

Hudson, Richard (2003a) Case-agreement, PRO and structure sharing. *Research in Language* (University of Lodz) 1: 7–33.

Hudson, Richard (2003b) Gerunds without phrase structure. *Natural Language and Linguistic Theory* 21: 579–615.

Hudson, Richard (2003c) Mismatches in default inheritance. In E. Francis and L. Michaelis (eds.), *Mismatch: Form-Function Incongruity and the Architecture of Grammar*, pp. 269–317. Stanford, CA: CSLI Publications.

Hudson, Richard (2003d) Trouble on the left periphery. *Lingua* 113: 607–42.

Hudson, Richard (2006)*Language Networks: The New Word Grammar*. Oxford: Oxford University Press.

Hudson, Richard (2007) English dialect syntax in Word Grammar. *English Language and Linguistics* 11: 383–405.

Hudson, Richard, and J. Holmes (2000) Re-cycling in the Encyclopedia. In Bert Peeters (ed.), *The Lexicon/Encyclopedia Interface*, pp. 259–90. Amsterdam: Elsevier.

Humboldt, Wilhelm von (1971[1836]) *Linguistic Variability* and *Intellectual Development* (English translation). Gainesville: University of Florida Press.

Humphreys, Glyn W., and Jane M. Riddoch (eds.) (1987) *Visual Object Processing: A Cognitive Neuropsychological Approach*. Mahwah, NJ: Erlbaum.

Hunston, S., and G. Francis (1998) Verbs observed: a corpus-driven pedagogic grammar. *Applied Linguistics* 19: 45–72.

Hunston, S., and G. Francis (2000) *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: Benjamins.

Huybregts, Riny (1984) The weak inadequacy of context-free phrase structure grammars. In Ger de Haan, M. Trommelen, and W. Zonneveld (eds.), *Van Periferie naar Kern*, pp. 81–99. Dordrecht: Foris.

Hyman, Larry M., and William R. Leben (2000) Suprasegmental units. In Geert Booij, Christian Lehmann, and Joachim Mugdan (eds.), *Morphology/Morphologie: A Handbook on Inflection and Word Formation/Ein Handbuch zur Flexion und Wortbildung*, vol. 1, pp. 587–94. Berlin: De Gruyter.

Ifantidou-Trouki, Elly (1993) Sentential adverbs and relevance. *Lingua* 90: 69–90.

Iggesen, Oliver A. (2005) Number of cases. In *WALS*, pp. 202–5.

Imai, Mutsumi, and Reiko Mazuka (2003) Reevaluating linguistic relativity: language-specific categories and the role of universal ontological knowledge in the construal of individuation. In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind*, pp. 429–64. Cambridge, MA: MIT Press.

Inkelas, Sharon (1993) Nimboran position class morphology. *Natural Language and Linguistic Theory* 11: 559–624.

Ito, Junko (1986) Syllable theory in prosodic phonology. Doctoral dissertation, University of Massachusetts, Amherst. Published 1988, Outstanding Dissertations in Linguistics series, New York: Garland.

Ito, Junko, and Armin Mester (1994) Reflections on CodaCond and Alignment. In Jason Merchant, Jaye Padgett, and Rachel Walker (eds.), *Phonology at Santa Cruz*, vol. 3, pp. 27–46. Santa Cruz: Linguistics Research Center, University of California, Santa Cruz.

Jackendoff, Ray (1977) *X Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1983) *Semantics and Cognition.* Cambridge, MA: MIT Press.

Jackendoff, Ray (1987) *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1990a) *Semantic Structures.* Cambridge: MIT Press.

Jackendoff, Ray (1990b) On Larson's treatment of the double object construction. *Linguistic Inquiry* 21: 427–6. http://en.wikipedia.org/wiki/Heavy_NP_shift - _ref-3

Jackendoff, Ray (1992) Mme. Tussaud meets the Binding Theory. *Natural Language and Linguistic Theory* 10: 1–31.

Jackendoff, Ray (1997) *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Jackendoff, Ray (2002) *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford: Oxford University Press.

Jackendoff, Ray (2007a) A parallel architecture perspective on language processing. *Brain Research* 1146: 2–22.

Jackendoff, Ray (2007b) *Language, Consciousness, Culture*. Cambridge, MA: MIT Press.

Jackendoff, Ray (2008) Alternative minimalist visions of language. In R. Edwards, P. Midtlying, K. Stensrud, and C. Sprague (eds.), *Chicago Linguistics Society 41: The Panels*. Also forthcoming in Robert D. Borsley and Kersti Börjars, *Nontransformational Syntax: A Guide to Current Models* (Oxford: Blackwell).

Jackendoff, Ray, and Peter W. Culicover (2003) The semantic basis of control. *Language* 79: 517–56.

Jackendoff, Ray, and Peter W. Culicover (2005) *Simpler Syntax*. Oxford: Oxford University Press.

Jackendoff, Ray, and Fred Lerdahl (2006) The capacity for music: what's special about it? *Cognition* 100: 33–72.

Jackendoff, Ray, and Steven Pinker (2005) The nature of the language faculty and its implications for the evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition* 97: 211–25.

Jacobs, Joachim (1994) *Kontra Valenz*. Trier: Wissenschaftlicher Verlag.

Jacobs, Joachim (2003) Die Problematik der Valenzebenen. In Ágel et al. (2003: 378–99).

Jacobs, Joachim (2008) Wozu Konstruktionen? *Linguistische Berichte* 213: 3–44.

Jacobsen, W. (1979) Noun and verb in Nootkan. In Barbara Efrat (ed.), *The Victoria Conference on Northwestern Languages*, pp. 83–155. Victoria: British Columbia Provincial Museum.

Jaeger, Florian, and Neil Snider (2008) Implicit learning and syntactic persistence: surprisal and cumulativity. *Proceedings CogSci 2008* (Washington, DC), pp. 1061–6.

Jäger, Gerhard, and Anette Rosenbach (2006) The winner takes it all—almost: cumulativity in grammatical variation. *Linguistics* 44: 937–71.

Jakobson, Roman (1941) *Kindersprache, Aphasie, und Allgemeine Lautgesetze*. Uppsala: Almqvist & Wiksell. Translated by A. Keiler as *Child Language, Aphasia and Phonological Universals*, Berlin: Mouton, 1968.

Jang, S. C. (1998) Dimensions of spoken and written Taiwanese: a corpus-based register study. PhD dissertation, University of Hawaii.

Janson, Tore (2004) *A Natural History of Latin*. Oxford: Oxford University Press.

January, David, and Edward Kako (2007) Re-evaluating evidence for linguistic relativity: reply to Boroditsky (2001). *Cognition* 104: 417–26.

Janzen, Terry (1995) The polygrammaticalization of FINISH in ASL. MS.

Janzen, Terry (1999) The grammaticization of topics in American Sign Language. *Studies in Language* 23: 271–306.

Janzen, Terry, and Barbara Shaffer (2002) Gesture as the substrate in the process of ASL grammaticization. In R. Meier, D. Quinto, and K. Cormier (eds.), *Modality and Structure in Signed and Spoken Languages*, pp. 199–223. Cambridge: Cambridge University Press.

Janzen, Terry, Barbara Shaffer, and Sherman Wilcox (2000) Signed language pragmatics. In Jef Verschueren, Jan-Ola Östman, Jan Blommaert, and Chris Bulcaen (eds.), *Handbook of Pragmatics*, p. 1–20. Amsterdam: Benjamins.

Jary, Mark (1998) Relevance theory and the communication of politeness. *Journal of Pragmatics* 30: 1–19.

Jaszczolt, Katarzyna M. (1992) Belief sentences and the semantics of propositional attitudes. DPhil. thesis, University of Oxford.

Jaszczolt, Katarzyna M. (1997) The default *de re* principle for the interpretation of belief utterances. *Journal of Pragmatics* 28: 315–36.

Jaszczolt, Katarzyna M. (1999) *Discourse, Beliefs, and Intentions: Semantic Defaults and Propositional Attitude Ascription*. Oxford: Elsevier Science.

Jaszczolt, Katarzyna M. (2005) *Default Semantics: Foundations of a Compositional Theory of Acts of Communication*. Oxford: Oxford University Press.

Jaszczolt, Katarzyna M. (2006) Defaults in semantics and pragmatics. In E. N. Zalta (ed.) *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/contents.html

Jaszczolt, Katarzyna M. (2007a) Variadic function and pragmatics-rich representations of belief reports. *Journal of Pragmatics* 39: 934–59.

Jaszczolt, Katarzyna M. (2007b) On being post-Gricean. In R. A. Nilsen, N. A. A. Amfo, and K. Borthen (eds.), *Interpreting Utterances: Pragmatics and Its Interfaces. Essays in Honour of Thorstein Fretheim*, pp. 21–38. Oslo: Novus.

Jaszczolt, Katarzyna M. (2008) Psychological explanations in Gricean pragmatics and Frege's legacy. In I. Kecskes and J. Mey (eds.), *Intentions, Common Ground, and the Egocentric Speaker-Hearer*, pp. 9–44. Berlin: Mouton de Gruyter.

Jaszczolt, Katarzyna M. (2009) *Representing Time: An Essay on Temporality as Modality*. Oxford: Oxford University Press.

Jaszczolt, Katarzyna M. (forthcoming a) Semantics and pragmatics: the boundary issue. In K. von Heusinger, P. Portner, and C. Maienborn (eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.

Jaszczolt, Katarzyna M. (forthcoming b) Semantics-pragmatics interface. In L. Cummings (ed.), *The Pragmatics Encyclopedia*. London: Routledge.

Jaszczolt, Katarzyna M., and Jiranthara Srioutai (forthcoming) Communicating about the past through modality in English and Thai. In F. Brisard and T. Mortelmans (eds.), *Cognitive Approaches to Tense, Aspect and Modality*. Amsterdam: Benjamins.

Jespersen, Otto (1924) *The Philosophy of Grammar*. New York: Norton (reprinted 1965).

Johnson, Christopher (1997) The acquisition of the 'What's X doing Y?' construction. In Mary Hughes and Annabel Greenhill (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, vol. 2, pp. 343–53. Somerville, MA: Cascadilla Press.

Johnson, David E., and Paul M. Postal (1980) *Arc Pair Grammar.* Princeton, NJ: Princeton University Press.

Johnson, Keith (1997) Speech perception without speaker normalization. In Keith Johnson and John W. Mullennix (eds.), *Talker Variability in Speech Processing*, pp. 145–65. San Diego, CA: Academic Press.

Johnson, Mark (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.

Johnson, Mark (2002) The DOP estimation method is biased and inconsistent. *Computational Linguistics* 28: 71–6.

Jolly, J. (1993) Preposition assignment in English. In Van Valin (1993a: 275–310).

Jones, Janet (2008) Multiliteracies for academic purposes: a metafunctional exploration of intersemiosis and multimodality in university textbook and computer-based learning resources in science. University of Sydney.

Joos, Martin (ed.) (1957) *Readings in Linguistics I.* Washington, DC: American Council of Learned Societies.

Joseph, Brian D. (1994) On weak subjects and pro-drop in Greek. In Irene Philippaki-Warburton, Katerina Nicolaidis, and Maria Sifianou (eds.), *Themes in Greek Linguistics (Papers from the First International Conference on Greek Linguistics, Reading, September 1993)*, pp. 21–32. Amsterdam: Benjamins.

Joseph, Brian D. (2001) Is there such a thing as grammaticalization? *Language Sciences* 23: 163–86.

Joseph, Brian D., and Richard D. Janda (eds.) (2003) *The Handbook of Historical Linguistics*. Oxford: Blackwell.

Joshi, Aravind (1987) An introduction to tree-adjoining grammars. In A. Manaster-Ramer (ed.), *Mathematics of Language*, pp. 87–114. Amsterdam: Benjamins.

Joshi, Aravind (2004) Starting with complex primitives pays off: complicate locally, simplify globally. *Cognitive Science* 28: 637–68.

Jucker, Andreas (1992) Pragmatics of the definite article in English. In F. J. H. Dols (ed.), *Pragmatic Grammar Components*, pp. 117–33. Tilburg: Tilburg University Press.

Jurafsky, Daniel (1996) A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20: 137–94.

Jurafsky, Daniel (2003) Probabilistic modeling in psycholinguistics. In R. Bod, J. Hay, and S. Jannedy (eds.), *Probabilistic Linguistics*, pp. 39–96. Cambridge, MA: MIT Press.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond (2001) Probabilistic relations between words: evidence from reduction in lexical production. In Bybee and Hopper (2001: 229–54).

Kaas, Jon H. (1989) Why does the brain have so many visual areas? *Journal of Cognitive Neuroscience* 1: 21–35.

Kager, René (1999) *Optimality Theory*. Cambridge: Cambridge University Press.

Kager, René (2001) Rhythmic directionality by positional licensing. Handout from HILP 5, University of Potsdam. Available on the Rutgers Optimality Archive, ROA 514, http://roa.rutgers.edu

Kager, René, Joe Pater, and Wim Zonneveld (2004) *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Kailuweit, R., T. Bender, M. Hartung, et al. (2003) Parsing English, and French using the RRG paradigm. Paper presented at the 2003 International Conference on Role and Reference Grammar, UNESP São José do Rio Preto, Brazil.

Kailuweit, R., E. Staudinger, B. Wiemer, and R. Matasović (eds.) (2008) *New Applications of Role and Reference Grammar: Diachrony, Grammaticalization, Romance Languages.*Newcastle: Cambridge Scholars Press.

Kam, X., I. Stoyneshka, L. Tornyova, W. Sakas, and J. Fodor (2005) Statistics vs. UG in language acquisition: does a bigram analysis predict auxiliary inversion? *Proceedings of the 2nd ACL Workshop on Psychocomputational Models of Human Language Acquisition*, pp. 69–71.

Kamp, Hans, and Uwe Reyle (1993) *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic.

Kamp, Hans, and Jan van Eijck (1996) Representing discourse in context. In J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Linguistics*, pp. 179–237. Amsterdam: Elsevier.

Kanoksilapatham, Budsaba (2007) Rhetorical moves in biochemistry research articles. In Biber, Connor, and Upton (2007: 73–120).

Kao, Rong-Rong (1993) Grammatical relations and anaphoric structures in Mandarin Chinese. PhD dissertation, University of Hawaiʻi at Manoa.

Kaplan, Ronald M. (1987) Three seductions of computational psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson, and Paul Bennett (eds.), *Linguistic Theory and Computer Applications*, pp. 149–81. London: Academic Press. Reprinted in Dalrymple et al. (1995: 339–67).

Kaplan, Ronald M. (1989) The formal architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen (eds.), *Proceedings of ROCLING II*, pp. 3–18. Reprinted in Dalrymple et al. (1995: 7–27).

Kaplan, Ronald M. (1996) A probabilistic approach to lexical-functional analysis. *Proceedings of the 1996 LFG Conference and Workshops*. Stanford, CA: CSLI Publications.

Kaplan, Ronald M., and Joan Bresnan (1982) Lexical-Functional Grammar: a formal system for grammatical representation. In Bresnan (1982a: 173–281). Reprinted in Dalrymple et al. (1995: 29–135).

Kaplan, Ronald M., and John T. Maxwell, III (1988) Constituent coordination in Lexical-Functional Grammar. *Proceedings of COLING-88* (Budapest), vol. 1. Reprinted in Dalrymple et al. (1995: 199–214).

Kaplan, Ronald M., and Annie Zaenen (1989) Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin and Anthony Kroch (eds.), *Alternative Conceptions of Phrase Structure*, pp. 17–42. Chicago: University of Chicago Press. Reprinted in Dalrymple et al. (1995: 137–65).

Karimi-Doostan, G. (1997) Light verb constructions in Persian and Kurdish. PhD dissertation, University of Essex.

Karimi, Simin (2003) On object positions, specificity and scrambling in Persian. In Simin Karimi (ed.), *Word Order and Scrambling*, pp. 91–125. Oxford: Blackwell.

Karlgren, J. (2000) Stylistic experiments for informational retrieval. PhD dissertation, Stockholm University.

Kaschak, Michael P., and Arthur M. Glenburg (2004) This construction needs learned. *Journal of Experimental Psychology: General* 133: 450–67.

Kasher, Asa (1976) Conversational maxims and rationality. In A. Kasher (ed.), *Language in Focus*, pp. 197–216. Dordrecht: Reidel.

Katz, Jerrold J., and Jerry A. Fodor (1963) The structure of a semantic theory. *Language* 39: 170–210.

Kauffman, Stuart A. (1993) *The Origins of Order*. Oxford: Oxford University Press.

Kawahara, Shigeto (2006) A faithfulness ranking projected from a perceptibility scale: the case of [+voice] in Japanese. *Language* 82: 536–74.

Kay, Martin (1980) *Algorithmic Schemata and Data Structures in Syntactic Processing*. Report CSL-80-12, Xerox PARC, Palo Alto, CA.

Kay, Paul, and Charles Fillmore (1999) Grammatical constructions and linguistic generalizations: the What's X doing Y? construction. *Language* 75: 1–33.

Kay, Paul, and Willett Kempton (1984) What is the Sapir–Whorf hypothesis? *American Anthropologist* 86: 65–79.

Kay, Paul, and Laura A. Michaelis (forthcoming) Constructions and compositionality. In C. Maienborn, K. von Heusinger, and P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.

Kay, Paul, and Terry Regier (2007) Color naming universals: the case of Berinmo. *Cognition* 102: 289–98.

Kayne, Richard S. (1984) *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, Richard S. (1994) *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, Richard S. (2003) Silent years, silent hours. In L.-O. Delsing, C. Falk, G. Josefsson, and H. Á. Sigurðsson (eds.), *Grammatik i fokus: Festskrift till Christer Platzack den 18 november 2003*, pp. 209–26. Lund: Wallin & Dalholm.

Kayne, Richard S. (2005a) Some notes on comparative syntax, with special reference to English and French. In G. Cinque and R. S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, pp. 3–69. New York: Oxford University Press.

Kayne, Richard S. (2005b) *Movement and Silence*. New York: Oxford University Press.

KAYNE, RICHARD S. (2006) On parameters and on principles of pronunciation. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz, and J. Koster (eds.), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, pp. 289–99. Berlin: Mouton de Gruyter.

KEENAN, EDWARD L. (1985) Relative clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 2:*Complex Constructions*, pp. 141–234. Cambridge: Cambridge University Press.

KEENAN, EDWARD L., and BERNARD COMRIE (1977) Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.

KEHLER, ANDREW (2002) *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI Publications.

KELLER, RUDY (1994) *On Language Change: The Invisible Hand in Language*. London: Routledge.

KELLY, MICHAEL H., and SUSANNE MARTIN (1994) Domain-general abilities applied to domain-specific tasks: sensitivity to probabilities in perception, cognition and language. *Lingua* 92: 105–40.

KEMPEN, GERARD, and ERIC HOENKAMP (1987) An incremental procedural grammar for sentence formation. *Cognitive Science* 11: 201–58.

KEMPSON, RUTH M. (1975) *Presupposition and the Delimitation of Semantics*. Cambridge: Cambridge University Press.

KEMPSON, RUTH M. (1979) Presupposition, opacity, and ambiguity. In C.-K. Oh and D. A. Dinneen (eds.), *Syntax and Semantics*, vol. 11, pp. 283–97. New York: Academic Press.

KEMPSON, RUTH M. (1980) Ambiguity and word meaning. In S. Greenbaum, G. Leech, and J. Svartvik (eds.), *Studies in English Linguistics*, pp. 7–16. London: Longman.

KENESEI, ISTVÁN, ROBERT M. VAGO, and ANNA FENYVESI (1998) *Hungarian*. London: Routledge.

KENNEDY, G. (1998) *An Introduction to Corpus Linguistics*. London: Longman.

KENSTOWICZ, MICHAEL (1996) Quality-sensitive stress. *Rivista di linguistica* 9: 157–87.

KIEFER, FERENC, and LÁSZLÓ HONTI (2003) Verbal 'prefixation' in the Uralic languages. *Acta Linguistica Hungarica* 50: 137–53.

KIM, YONGJIN (1990) Register variation in Korean: a corpus-based study. Doctoral dissertation, University of Southern California.

KIM, YONGJIN, and DOUGLAS BIBER (1994) A corpus-based analysis of register variation in Korean. In Douglas Biber and E. Finegan (eds.), *Sociolinguistic Perspectives on Register*, pp. 157–81. New York: Oxford University Press.

KING, JEFFREY and JASON STANLEY (2005) Semantics, pragmatics, and the role of semantic content. In Z. G. Szabó (ed.), *Semantics vs. Pragmatics*, pp. 111–64. Oxford: Clarendon Press.

KING, TRACY HOLLOWAY (1995) *Configuring Topic and Focus in Russian*. Stanford, CA: CSLI Publications.

KINGSTON, JOHN (1999) Extrapolating from spoken to signed prosody via laboratory phonology. *Language and Speech* 42: 251–81.

KINTSCH, WALTER (1992) How readers construct situation models for stories: the role of syntactic cues and causal inference. In A. F. Healy, S. Kosslyn, and R. M. Shiffrin (eds.), *Essays in Honor of William K. Estes*, vol. 2, pp. 261–78. Hillsdale, NJ: Erlbaum.

KINTSCH, WALTER, and TEUN VAN DIJK (1978) Toward a model of text comprehension and production. *Psychological Review* 85: 363–94.

Kiparsky, Paul (1982) Lexical morphology and phonology. In I.-S. Yang (ed.), *Linguistics in the Morning Calm*, vol. 1, pp. 3–91. Seoul: Hanshin.

Kiparsky, Paul (1983) Word-formation and the lexicon. *Proceedings of the 1982 Mid-America Linguistics Conference*, pp. 47–78. Lawrence: University of Kansas, Dept of Linguistics.

Kiparsky, Paul (1985) Some consequences of lexical phonology. *Phonology Yearbook* 2: 85–138.

Kiparsky, Paul (1993) Variable rules. Paper presented at Rutgers Optimality Workshop I.

Kiparsky, Paul (2005) Grammaticalization as optimization. TS, Stanford University.

Kiparsky, Paul (forthcoming) *Paradigm Effects and Opacity*. Stanford, CA: CSLI Publications.

Kirby, Simon (1999) *Function, Selection, and Innateness: The Emergence of Language Universals.* Oxford: Oxford University Press.

Kirchner, Robert (1993) Turkish vowel harmony and disharmony: an Optimality Theoretic account. MS. Rutgers Optimality Archive, ROA-4.

Kirchner, Robert (1998) An effort-based approach to consonant lenition. Doctoral dissertation, University of California, Los Angeles.

Kiss, Katalin (ed.) (1995) *Discourse-Configurational Languages*. New York: Oxford University Press.

Kisseberth, Charles (1970a) On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.

Kisseberth, Charles (1970b) Vowel elision in Tonkawa and derivational constraints. In Jerrold M. Sadock and Anthony L. Vanek (eds.), *Studies Presented to Robert B. Lees by his Students*, pp. 109–37. Edmonton, Alberta: Linguistic Research.

Kleene, Stephen C. (1956) Representation of events in nerve nets and finite automata. In Claude Shannon and John McCarthy (eds.), *Automata Studies*, pp. 3–41. Princeton, NJ: Princeton University Press.

Klein, Dan, and Chris Manning (2003) Accurate unlexicalized parsing. *Proceedings ACL 2003*, pp. 423–30.

Klein, Dan, and Chris Manning (2004) Corpus-based induction of syntactic structure: models of dependency and constituency. *Proceedings ACL 2004* (Barcelona).

Klein, Dan, and Chris Manning (2005) Natural language grammar induction with a generative constituent-context model. *Pattern Recognition* 38: 1407–19.

Klima, Edward, and Ursula Bellugi (1979) *The Signs of Language*. Cambridge, MA: Harvard University Press.

Kluender, Robert (1998) On the distinction between strong and weak islands: a processing perspective. In P. Culicover and L. McNally (eds.), *The Limits of Syntax*, pp. 241–79. San Diego, CA: Academic Press.

Kluender, Robert, and Marta Kutas (1993) Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8: 573–633.

Kobayashi, Ichiro, Moon-Soo Chang, and Michio Sugeno (2002) A study on meaning processing of dialogue with an example of development of travel consultation system. *Information Science* 144: 45–74.

Koch, Christof (2004) *The Quest for Consciousness*. Englewood, CO: Roberts.

Kodytek, Vilem (2007) On the replicability of the Biber model: the case of Czech. TS.

Koenig, Jean-Pierre (ed.) (1998) *Discourse and Cognition: Bridging the Gap*. Stanford, CA: CSLI Publications.

Kokkonidis, Miltiadis (2008) First-order glue. *Journal of Logic, Language and Information* 17: 43–68.

König, Ekkehard, and Volker Gast (eds.) (forthcoming) *Reciprocals and Reflexives: Theoretical and Typological Explorations*. Berlin: Mouton de Gruyter.

Koopman, Hilda, and Dominique Sportiche (1991) The position of subjects. *Lingua* 85: 211–58.

Koopman, Hilda, and Anna Szabolcsi (2000) *Verbal Complexes*. Cambridge, MA: MIT Press.

Korhonen, Jarmo (2006) Valenzwandel am Beispiel des Deutschen. In Ágel et al. (2006: 1462–74).

Kortmann, Bernd, and Ekkehard König (1992) Categorial reanalysis: the case of deverbal prepositions. *Linguistics* 30: 671–97.

Koster, Jan (1978) *Locality Principles in Syntax*. Dordrecht: Foris.

Kövecses, Zoltán (2005) *Metaphor in Culture: Universality and Variation*. Cambridge: Cambridge University Press.

Kozinsky, Ivan. S., Vladimir P. Nedjalkov, and Maria S. Polinskaja (1988) Antipassive in Chukchee: oblique object, object incorporation, zero object. In Masayoshi Shibatani (ed.), *Passive and Voice*, pp. 651–706. Amsterdam: Benjamins.

Kress, Gunther, and Theo van Leeuwen (2006) *Reading Images*, 2nd edn. New York: Routledge.

Krifka, Manfred (1995) Common nouns: a contrastive analysis of Chinese and English. In Greg Carlson and Francis Pelletier (eds.), *The Generic Book*, pp. 398–411. Chicago: University of Chicago Press.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory N. Carlson, Alice ter Meulen, Godehard Link, and Gennaro Chierchia (1995) Genericity: an introduction. In Greg Carlson and Francis Pelletier (eds.), *The Generic Book*, pp. 1–124. Chicago: University of Chicago Press.

Kroeger, Paul (1993) *Phrase Structure and Grammatical Relations in Tagalog*. Stanford, CA: CSLI Publications.

Kroeger, Paul (2004) *Analyzing Syntax: A Lexical-Functional Approach*. Cambridge: Cambridge University Press.

Kroeger, Paul (2007) LFG as a framework for descriptive grammar. In Butt and King (2007), pp. 293–301.

Krug, Manfred (2000) *Emerging English Modals: A Corpus-Based Study of Grammaticalization*. Berlin: Mouton de Gruyter.

Krug, Manfred (2003) Frequency as a determinant in grammatical variation and change. In Rohdenburg and Mondorf (2003: 7–67).

Kuhn, Jonas (2003) *Optimality-Theoretic Syntax: A Declarative Approach*. Stanford, CA: CSLI Publications.

Kuno, Susumu, and Ken-Ichi Takami (1993) *Grammar and Discourse Principles*. Chicago: University of Chicago Press.

Kury-owicz, Jerzy (1965) The evolution of grammatical categories. *Diogenes* 51: 55–71.

Kuteva, Tania (1998) On identifying an evasive gram: action narrowly averted. *Studies in Language* 22: 113–60.

Kuteva, Tania (1999) On 'sit'/'stand'/'lie' auxiliation. *Linguistics* 37: 191–213.

Kuteva, Tania (2001) *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.

Labov, William (1966) *The Social Stratificiation of English in New York City*. Washington DC: Center for Applied Linguistics.

Labov, William (1969) Contraction, deletion, and inherent variability of the English copula. *Language* 45: 715–62.

Labov, William (1972) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William (1973) The boundaries of words and their meanings. In Charles-James Bailey and Roger Shuy (eds.), *New Ways of Analyzing Variation in English*, pp. 340–73. Washington, DC: Georgetown University Press.

Ladusaw, William, and David Dowty (1988) Toward a nongrammatical account of thematic roles. In Wendy Wilkins (ed.), *Thematic Relations*, pp. 62–74. San Diego, CA: Academic Press.

Lakoff, George (1970) *Irregularity in Syntax*. New York: Holt, Rinehart & Winston.

Lakoff, George (1971) On generative semantics. In D. Steinberg and L. Jakobovits (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, pp. 232–96. New York: Cambridge University Press.

Lakoff, George (1977) Linguistic gestalts. *Papers From the Regional Meeting of the Chicago Linguistic Society* 13: 236–87.

Lakoff, George (1987) *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: University of Chicago Press.

Lakoff, George (2008) *The Political Mind*. New York: Penguin.

Lakoff, George, and Mark Johnson (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, George, and Mark Johnson (1999) *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lakoff, George, and Rafael E. Núñez (2000) *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.

Lakoff, Robin (1995) Conversational logic. In J. Verschueren, Jan-Ola Östman, and Jan Blommaert (eds.), *Handbook of Pragmatics*, pp. 190–98. Amsterdam: Benjamins.

Lamb, Sydney (1966) *Outline of Stratificational Grammar*. Washington, DC: Georgetown University Press.

Lamb, Sidney (1998) *Pathways of the Brain: The Neurocognitive Basis of Language.* Amsterdam: Benjamins.

Lambek, Joachim (1988) The mathematics of sentence structure. *American Mathematical Monthly* 65: 154–70. Reprinted in Wojciech Buszkowski, Wojciech Marciszewski, and Johan van Benthem (eds.), *Categorial Grammar*, pp. 153–72. Amsterdam: Benjamins.

Lambrecht, Knud (1990) What me worry? *Mad Magazine* sentences revisited. *Berkeley Linguistics Society* 16: 215–28.

Lambrecht, Knud (1994) *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

Landau, Ilan (2000) *Elements of Control*. Dordrecht: Kluwer Academic.

Lane, Harlan (1984) *When the Mind Hears: A History of the Deaf*. New York: Random House.

Langacker, Ronald W. (1976) Semantic representations and the linguistic relativity hypothesis. *Foundations of Language* 14: 307–57.

Langacker, Ronald W. (1982) Space grammar, analysability, and the English passive. *Language* 58: 22–80.

LANGACKER, RONALD W. (1987a) Nouns and verbs. *Language* 63: 53–94.

LANGACKER, RONALD W. (1987b) *Foundations of Cognitive Grammar*, vol. 1: *Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

LANGACKER, RONALD W. (1987c) An introduction to cognitive grammar. *Cognitive Science* 10: 1–40.

LANGACKER, RONALD W. (1988) A usage-based model. In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, pp. 127–61. Amsterdam: Benjamins.

LANGACKER, RONALD W. (1990) *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.

LANGACKER, RONALD W. (1991) *Foundations of Cognitive Grammar*, vol. 2: *Descriptive Applications*. Stanford, CA: Stanford University Press.

LANGACKER, RONALD W. (1993) Grammatical traces of some invisible semantic constructs. *Language Sciences* 15: 323–55.

LANGACKER, RONALD W. (1995a) Structural syntax: the view from Cognitive Grammar. In Françoise Madray-Lesigne and Jeannine Richard-Zappella (eds.), *Lucien Tesnière Aujourd'hui*, pp. 13–39. Paris: Peeters.

LANGACKER, RONALD W. (1995b) Raising and transparency. *Language* 71: 1–62.

LANGACKER, RONALD W. (1997) Constituency, dependency, and conceptual grouping. *Cognitive Linguistics* 8: 1–32.

LANGACKER, RONALD W. (1998a) On subjectification and grammaticalization. In Koenig (1998: 71–89).

LANGACKER, RONALD W. (1998b) Conceptualization, symbolization and grammar. In Michael Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, pp. 1–39. Mahwah, NJ: Erlbaum.

LANGACKER, RONALD W. (1999a) Assessing the cognitive linguistic enterprise. In Theo Janssen and Gisela Redeker (eds.), *Cognitive Linguistics: Foundations, Scope, and Methodology*, pp. 13–59. Berlin: Mouton de Gruyter.

LANGACKER, RONALD W. (1999b) *Grammar and Conceptualization*. Berlin: Mouton de Gruyter.

LANGACKER, RONALD W. (1999c) Virtual reality. *Studies in the Linguistic Sciences* 29: 77–103.

LANGACKER, RONALD W. (2000) A dynamic usage-based model. In Michael Barlow and Suzanne Kemmer (eds.), *Usage-Based Models of Language*, pp. 1–63. Stanford, CA: CSLI Publications.

LANGACKER, RONALD W. (2004) Form, meaning, and behavior: the Cognitive Grammar analysis of double subject constructions. In Ellen Contini-Morava, Robert S. Kirsner, and Betsy Rodríguez-Bachiller (eds.), *Cognitive and Communicative Approaches to Linguistic Analysis*, pp. 21–60. Amsterdam: Benjamins.

LANGACKER, RONALD W. (2005a) Construction Grammars: cognitive, radical, and less so. In Francisco J. Ruiz de Mendoza Ibáñez and M. Sandra Peña Cervel (eds.), *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*, pp. 101–59. Berlin: Mouton de Gruyter.

LANGACKER, RONALD W. (2005b) Integration, grammaticization, and constructional meaning. In Mirjam Fried and Hans C. Boas (eds.), *Grammatical Constructions: Back to the Roots*, pp. 157–89. Amsterdam: Benjamins.

LANGACKER, RONALD W. (2006) On the continuous debate about discreteness. *Cognitive Linguistics* 17: 107–51.

Langacker, Ronald W. (2007) Constructing the meanings of personal pronouns. In Günter Radden, Klaus-Michael Köpke, Thomas Berg, and Peter Siemund (eds.), *Aspects of Meaning Construction*, pp. 171–87. Amsterdam: Benjamins.

Langacker, Ronald W. (2008) *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.

LaPolla, Randy, and Dory Poa (2002) On describing word order. TS, City University of Hong Kong (downloadable from http://personal.cityu.edu.hk/%7Ectrandy/describingwo.pdf)

Lardiere, Donna (1992) On the linguistic shaping of thought: another response to Alfred Bloom. *Language in Society* 21: 231–51.

Larsen-Freeman, Diane (1997) Chaos/complexity science and second language acquisition. *Applied Linguistics* 18: 141–65.

Larson, Richard K. (1988) On the double object construction. *Linguistic Inquiry* 19: 335–92.

Lasnik, Howard, and Mamoru Saito (1984) On the nature of proper government. *Linguistic Inquiry* 15: 235–89.

Lasnik, Howard, and Mamoru Saito (1992) *Move α*. Cambridge, MA: MIT Press.

Lasnik, Howard, and Juan Uriagereka (1988) *Lectures in GB Syntax*. Cambridge, MA: MIT Press.

Lasnik, Howard, Juan Uriagereka, and Cedric Boeckx (2005) *A Course in Minimalist Syntax*. Oxford: Blackwell.

László, Sarolta (1988) Mikroebene. In Pavica Mrazović and Wolfgang Teubert (eds.), *Valenzen im Kontrast: Ulrich Engel zum 60. Geburtstag*, pp. 218–33. Heidelberg: Groos.

Lazard, Gilbert (2006) *La Quête des invariants interlangues: la linguistique est-elle une science?* Paris: Champion.

Lee, Chungmin (1999) Topic, contrastive topic and focus: what's on our minds. Available at http://www.mind.sccs.chukyo-u.ac.jp/jcss/ICCS/99/olp/pt1/pt1.htm

Lee, David (2000) Modelling variation in spoken and written English: the multi-dimensional approach revisited. PhD dissertation, Lancaster University.

Lee, Hanjung (2001) Optimization in argument expression and interpretation: a unified approach. PhD thesis, Stanford University.

Lee, Hye-Kyung (2002) The semantics and pragmatics of connectives with reference to English and Korean. PhD dissertation, University of Cambridge.

Lee, Penny (1996) *The Whorf Theory Complex: A Critical Reconstruction*. Amsterdam: Benjamins.

Lee, Penny (2000) When is 'linguistic relativity' Whorf's linguistic relativity? In M. Pütz and M. Verspoor (eds.), *Explorations in Linguistic Relativity*, pp. 45–68. Amsterdam: Benjamins.

Legate, Julie (2008) Morphological and abstract case. *Linguisitic Inquiry* 39: 55–102.

Legendre, Géraldine, Jane Grimshaw, and Sten Vikner (eds.) *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press.

Lehmann, Christian (1982[1995]) *Thoughts on Grammaticalization*. Munich: Lincom Europa.

Lehmann, Christian (1985) Grammaticalization: synchronic variation and diachronic change. *Lingua e stile* 20: 303–18.

Lehmann, Christian (2004) Interlinear morphemic glossing. In Geert Booij, Christian Lehmann, and Joachim Mugdan (eds.), *Morphology/Morphologie: A Handbook* *on Inflection and Word Formation/Ein Handbuch zur Flexion und Wortbildung*, vol. 2, pp. 1834–57. Berlin: de Gruyter.

Leiss, Elisabeth (2002) Die Wortart 'Verb'. In D. Alan Cruise, Frans Hundsnurscher, Michael Job, and Peter Lutzeier (eds.), *Lexikologie: Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*, vol. 1, pp. 605–16. Berlin: De Gruyter.

Leiss, Elisabeth (2003) Empirische Argumente für Dependenz. In Ágel et al. (2003: 311–24).

Lenneberg, Eric H. (1967) *Biological Foundations of Language*. New York: Wiley.

Lenneberg, Eric H., and John M. Roberts (1956) *The Language of Experience: A Study in Methodology*. *International Journal of American Linguistics*, Memoir 13. Bloomington: Indiana University.

Lerdahl, Fred, and Ray Jackendoff (1983) *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Leslie, Alan M., and Uta Frith (1988) Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology* 6: 315–24.

Lev, Iddo (2007) Packed computation of exact meaning representations. PhD thesis, Stanford University.

Levelt, Willem J. M. (1989) *Speaking*. Cambridge, MA: MIT Press.

Levelt, Willem J. M., Ardi Roelofs, and Antje Meyer (1999) A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22:–1–45.

Levin, Beth, and Malka Rappaport Hovav (1995) *Unaccusativity*. Cambridge, MA: MIT Press.

Levin, Beth, and Malka Rappaport Hovav (2005) *Argument Realization*. Cambridge: Cambridge University Press.

Levin, Juliette (1985) A metrical theory of syllabicity. Doctoral dissertation, MIT.

Levin, Juliette (1988) Generating ternary feet. *Texas Linguistic Forum* 29: 97–113.

Levin, Lori S. (1985) Operations on lexical forms: unaccusative rules in Germanic languages. PhD thesis, MIT. Reprinted as Levin (1988).

Levin, Lori S. (1988) *Operations on Lexical Forms: Unaccusative Rules in Germanic Languages*. New York: Garland.

Levinson, Stephen C. (1983) *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, Stephen C. (1987a) Minimization and conversational inference. In J. Verschueren and M. Bertuccelli-Papi (eds.) *The Pragmatic Perspective: Selected Papers from the 1985 International Pragmatics Conference*, pp. 61–129. Amsterdam: Benjamins.

Levinson, Stephen C. (1987b) Pragmatics and the grammar of anaphora: a partial pragmatic reduction of Binding and Control phenomena. *Journal of Linguistics* 23: 379–434.

Levinson, Stephen C. (1991) Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics* 27: 107–61.

Levinson, Stephen C. (1996a) Frames of references and Molyneux's question: crosslinguistic evidence. In P. Bloom, M. Peterson, L. Nadel, and M. Garrett (eds.), *Language and Space*, pp. 109–69. Cambridge, MA: MIT Press.

Levinson, Stephen C. (1996b) Relativity in spatial conception and description. In John Gumperz and Stephen Levinson (eds.), *Rethinking Linguistic Relativity*, pp. 177–202. Cambridge: Cambridge University Press.

Levinson, Stephen C. (1997) From outer to inner space: linguistic categories and non-linguistic thinking. In Jan Nuyts and Eric Pederson (eds.), *Language and Conceptualization*, pp. 13–45. Cambridge: Cambridge University Press.

Levinson, Stephen C. (2000) *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Levinson, Stephen C. (2003a) Language and the mind: let's get the issues straight! In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind*, pp. 25–46. Cambridge, MA: MIT Press. Levinson, Stephen C. (2003b) *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.

Levinson, Stephen C., Kita Sotaro, Daniel B. M. Hauna, and Björn H. Rasch (2002) Returning the tables: language affects spatial reasoning. *Cognition* 84: 155–88.

Levy, Roger (2008) Expectation-based syntactic comprehension. *Cognition* 106: 1126–77.

Lewis, John, and Jeffrey Elman (2001) Learnability and the statistical structure of language: poverty of stimulus arguments revisited. *Proceedings of 26th annual Boston University Conference on Language Development*, pp. 359–70.

Lewontin, Richard (1990) The evolution of cognition: questions we will never answer. In Daniel N. Osherson (ed.), *An Invitation to Cognitive Sciences*, vol. 4, pp. 107–32. Cambridge, MA: MIT Press.

Lewontin, Richard (2000) *The Triple Helix*. Cambridge, MA: Harvard University Press.

Li, Charles N. (ed.) (1975) *Word Order and Word Order Change.* Austin: University of Texas Press.

Li, Charles N. (2002) Missing links, issues and hypotheses in the evolutionary origin of language. In Givón and Malle (2002: 83–106).

Li, Charles N., and Sandra A. Thompson (1976) Subject and topic: a new typology of language. In Charles N. Li (ed.), *Subject and Topic*, pp. 459–89. New York: Academic Press.

Li, Charles N., and Sandra A. Thompson (1981) *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Li, Fei Fei, Rufin VanRullen, Christof Koch, and Pietro Perona (2002) Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences* 99: 9596–9601.

Li, Peggy, and Lila Gleitman (2002) Turning the tables: language and spatial reasoning. *Cognition* 83: 265–94.

Liddell, Scott K. (1984) THINK and BELIEVE: sequentiality in American Sign Language. *Language* 60: 372–99.

Liddell, Scott K. (2003) *Grammar, Gesture, and Meaning in American Sign Language.* New York: Cambridge University Press.

Lieberman, P. (1984) *The Biology and Evolution of Language*. Cambridge, MA: Harvard University Press.

Lieven, Elena, Heike Behrens, Jenny Speares, and Michael Tomasello (2003) Early syntactic creativity: a usage-based approach. *Journal of Child Language* 30: 333–70.

Lieven, Elena, Julian Pine, and Gillian Baldwin (1997) Lexically-based learning and early grammatical development. *Journal of Child Language* 24: 187–219.

Lightfoot, David (2006) *How New Languages Emerge*. Cambridge: Cambridge University Press.

Lindblom, Björn, Peter MacNeilage, and Michael Studdert-Kennedy (1984) Self-organizing processes and the explanation of phonological universals. In Brian Butterworth, Bernard Comrie, and Östen Dahl (eds.), *Explanations for Language Universals*, pp. 181–203. New York: Mouton.

Lindquist, H., and C. Mair (2004) *Corpus Approaches to Grammaticalization in English*. Amsterdam: Benjamins.

Lindquist, Kristen A., Lisa Feldman Barrett, Eliza Bliss-Moreau, and James A. Russell (2006) Language and the perception of emotion. *Emotion* 6: 125–38.

Liu, Haitao, Richard Hudson, and Zhiwei Feng (2008) Using a Chinese treebank to measure dependency distance. *Corpus Linguistics and Linguistic Theory*.

Liu, Lisa Gabern (1985) Reasoning counterfactually in Chinese: are there any obstacles? *Cognition* 21: 239–70.

Lobin, Henning (1993) *Koordinationssyntax als prozedurales Phänomen*. Tübingen: Narr.

Loftus, Elizabeth F. (1975) Leading questions and the eyewitness report. *Cognitive Psychology* 7: 560–72.

Loftus, E. F., and J. C. Palmer (1974) Reconstruction of automobile destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior* 13: 585–9.

Lombardi, Linda (ed.) (2001a) *Segmental Phonology in Optimality Theory: Constraints and Representations*. Cambridge: Cambridge University Press.

Lombardi, Linda (2001b) Why Place and Voice are different: constraint-specific alternations in Optimality Theory. In Lombardi (2001a: 13–45). *Longman Dictionary of Contemporary English* (1995) 3rd edn. Harlow: Longman.

Longobardi, Giuseppe (2001) How comparative is semantics– A unified parametric theory of bare nouns and proper names. *Natural Language Semantics* 9: 335–69.

Lönneker-Rodman, Birte (2007) Multilinguality and FrameNet. Technical Report TR-07-001, ICSI, Berkeley, CA.

Loucks, Jeff, and Eric Pederson (forthcoming) Linguistic and non-linguistic categorization of complex motion events. In J. Bohnemeyer and E. Pederson (eds.), *Event Representation in Language*. Cambridge: Cambridge University Press.

Lucy, John A. (1992a) *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. New York: Cambridge University Press.

Lucy, John A. (1992b) Language diversity and thought: a reformulation of the linguistic relativity hypothesis. New York: Cambridge University Press.

Lucy, John A. (1996) The scope of linguistic relativity: an analysis and review of empirical research. In J. Gumperz and S. Levinson (eds.), *Rethinking Linguistic Relativity*, pp. 37–69. Cambridge: Cambridge University Press.

Lucy, John A. (1997) Linguistic relativity. *Annual Review of Anthropology* 26: 291–312.

Lucy, John A., and Suzanne Gaskins (2001) Grammatical categories and the development of classification preferences: a comparative approach. In M. Bowerman and S. Levinson (eds.), *Language Acquisition and Conceptual Development*, pp. 257–83. Cambridge: Cambridge University Press.

Lucy, John A., and Suzanne Gaskins (2003) Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind*, pp. 465–92. Cambridge, MA: MIT Press.

Luger, George, and William Stubblefield (1993) *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. New York: Cummings.

Lynch, Michael (2007) *The Origins of Genome Architecture*. Sunderland, MA: Sinauer.

MacWhinney, Brian (1978) *The Acquisition of Morphophonology*. Chicago: University of Chicago Press.

MacWhinney, Brian (2000) *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Erlbaum.

MacWhinney, Brian (2005) Item-based constructions and the logical problem. *Proceedings of the Second Workshop on Psychocomputational Models of Human Language Acquisition* (Ann Arbor).

Mahlberg, M. (2005) *English General Nouns: A Corpus Theoretical Approach*. Amsterdam: Benjamins.

Mair, Christian (1990) *Infinitival Complement Clauses in English.* New York: Cambridge University Press.

Mairal Usón, Ricardo, and Pamela Faber (2002) Functional grammar and lexical templates. In R. Mairal Usón and M. J. Pérez Quintero (eds.), *New Perspectives on Argument Structure in Functional Grammar*, pp. 39–94. Berlin: Mouton de Gruyter.

Majid, Asifa (2002) Frames of reference and language concepts. *Trends in Cognitive Sciences* 6: 503–4.

Majid, Asifa, and Melissa Bowerman (eds.) (2007) Cutting and breaking events: a cross-linguistic perspective. Special issue of *Cognitive Linguistics* 18.2.

Majid, Asifa, Melissa Bowerman, Sotaro Kita, Daniel B. M. Haun, and Stephen C. Levinson (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8: 108–14.

Malle, Bertram F., Louis J. Moses, and Dare A. Baldwin (eds.) (2001) *Intentions and Intentionality: Foundations of Social Cognition*. Cambridge, MA: MIT Press.

Malt, Barbara C., and Edward E. Smith (1984) Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior* 23: 250–69.

Mandel, Mark A. (1977) Iconic devices in American Sign Language. In Lynn A. Friedman (ed.), *On the Other Hand: New Perspectives on American Sign Language*, pp. 57–108. New York: Academic Press.

Manning, Chris (2003) *Probabilistic Syntax*. In R. Bod, J. Hay, and S. Jannedy (eds.), *Probabilistic Linguistics*. Cambridge, MA: MIT Press, pp. 289–341.

Manning, Chris, and Hinrich Schütze (1999) *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Manzini, Maria Rita (1983) On Control and Control Theory. *Linguistic Inquiry* 14: 421–46.

Manzini, Maria Rita, and Leonardo Savoia (2005) *I dialetti italiani e romanzi*. Alessandria: Edizioni dell'Orso.

Marantz, Alec (1984) *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Marantz, Alec (1997) No escape from syntax. In A. Dimitriadis and I. Siegel (eds.), *University of Pennsylvania Working Papers in Linguistics*, pp. 201–25. Philadelphia: University of Pennsylvania Dept of Linguistics.

Marantz, Alec (2000) Words. MS, MIT.

Marantz, Alec (2005) Generative linguistics within the cognitive neuroscience of language.*Linguistic Review* 22: 429–45.

Marantz, Alec (2006) Phases and words, MS, NYU.

Marcus, Gary (2008) *Kluge*. New York: Houghton Mifflin.

Marcus, Mitchell (1980) *A Theory of Syntactic Recognition for Natural Language*. Cambridge, MA: MIT Press.

Marmaridou, Sophia (2006) On the conceptual, cultural and discursive motivation of Greek pain lexicalizations. *Cognitive Linguistics* 17: 393–434.

Martin, James R. (1990) Intrinsic functionality: implications for contextual theory. *Social Semiotics* 1: 99–162.

Martín Arista, J. (2008) Unification and separation in a functional theory of morphology. In Van Valin (2008b: 119–45).

Martinet, André (1962) *A Functional View of Language*. Oxford: Clarendon Press.

Martinet, André (1964) *Elements of General Linguistics*. London: Faber & Faber.

Martinich A. P. (1984) *Communication and Reference*. Berlin: de Gruyter.

Maslova, Elena, and Vladimir P. Nedjalkov (2005) Reciprocal constructions. In *WALS*, pp. 430–33.

Matasović, R. (2002) On representing syntactic change: towards a theory of diachronic syntax. *Suvremena lingvistika* 53–554: 57–72.

Matlock, Teenie (2004) The conceptual motivation of fictive motion. In Günter Radden and Klaus-Uwe Panther (eds.), *Studies in Linguistic Motivation*, pp. 221–48. Berlin: Mouton de Gruyter.

Matlock, Teenie, Michael Ramscar, and Lera Boroditsky (2005) The experiential link between spatial and temporal language. *Cognitive Science* 29: 655–64.

Matsumoto, Yo (1997) Linguistic evidence for subjective (fictive) motion. In Kei Yamanaka and Toshio Ohori (eds.), *The Locus of Meaning: Papers in Honor of Yoshihiko Ikegami*, pp. 209–20. Tokyo: Kuroshio.

Matthews, Peter H. (1981) *Syntax*. Cambridge: Cambridge University Press.

Matthews, Peter H. (2007) *Syntactic Relations: A Critical Survey.* Cambridge: Cambridge University Press.

Matthewson, Lisa (1998) *Determiner Systems and Quantificational Strategies: Evidence from Salish*. The Hague: Holland Academic Graphics.

Matthiessen, Christian M. I. M. (1993) Register in the round. In Mohsen Ghadessy (ed.), *Register Analysis: Theory and Practice*, pp. 221–92. London: Pinter.

Matthiessen, Christian M. I. M. (1995) *Lexicogrammatical Cartography: English Systems*. Tokyo: International Language Sciences Publishers.

Mattissen, Johanna (2006) The ontology and diachrony of polysynthesis. In Dieter Wunderlich (ed.), *Advances in the Theory of the Lexicon*, pp. 287–353. Berlin: Mouton de Gruyter. Maxwell, John T., III, and Ronald M. Kaplan (1991) A method for disjunctive constraint satisfaction. In Masaru Tomita (ed.), *Current Issues in Parsing Technology*, pp. 173–90. Dordrecht: Kluwer. Reprinted in Dalrymple et al. (1995: 381–401).

Maxwell, John T., III, and Ronald M. Kaplan (1993) The interface between phrasal and functional constraints. *Computational Linguistics* 19: 571–90. Reprinted in Dalrymple et al. (1995: 403–29).

Maxwell, John T., III, and Ronald M. Kaplan (1996) Unification-based parsers that automatically take advantage of context freeness. In Butt and King (1996).

Maxwell, John T., III, and Christopher D. Manning (1996) A theory of non-constituent coordination based on finite-state rules. In Butt and King (1996).

Mayr, Ernst (2004) *What Makes Biology Unique? Considerations on the Autonomy of a Scientific Discipline.* Cambridge: Cambridge University Press.

McCarthy, John J. (1986) OCP Effects: gemination and antigemination. *Linguistic Inquiry* 17: 207–63.

McCarthy, John J. (2002) *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. (ed.) (2003a) *Optimality Theory: A Reader*. Oxford: Blackwell.

McCarthy, John J. (2003b) OT constraints are categorical. *Phonology* 20: 75–138.

McCarthy, John J. (2003c) Sympathy, cumulativity, and the Duke-of-York gambit. In Caroline Fery and Ruben van de Vijver (eds.), *The Syllable in Optimality Theory*, pp. 23–76. Cambridge: Cambridge University Press.

McCarthy, John J. (2007a) *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.

McCarthy, John J. (2007b) What is Optimality Theory? *Language and Linguistics Compass* 10: 260–91.

McCarthy, John J. (2008) *Doing Optimality Theory*. Oxford: Blackwell.

McCarthy, John J., and Alan Prince (1993a) Generalized alignment. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology*, pp. 79–153. Dordrecht: Kluwer. Excerpts reprinted in John Goldsmith (ed.), *Essential Readings in Phonology* (Oxford: Blackwell, 1999), pp. 102–36.

McCarthy, John J., and Alan Prince (1993b) *Prosodic Morphology I: Constraint Interaction and Satisfaction*. New Brunswick, NJ: Rutgers University Center for Cognitive Science. Available on the Rutgers Optimality Archive, ROA 482, http://roa.rutgers.edu

McCarthy, John J., and Alan Prince (1994) The emergence of the unmarked: Optimality in prosodic morphology. In Merce Gonzalez (ed.), *Proceedings of the North East Linguistic Society 24*, pp. 333–79. Amherst, MA: GLSA Publications.

McCarthy, John J., and Alan Prince (1995) Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics 1*, pp. 249–384. Amherst, MA: GLSA Publications.

McCarthy, John J., and Matthew Wolf (2005) Less than zero: correspondence and the Null Output. MS. Available on the Rutgers Optimality Archive, ROA 722, http://roa.rutgers.edu

McCawley, James D. (1978) Conversational implicature and the lexicon. In P. Cole (ed.), *Pragmatics*, pp. 245–59. New York: Academic Press.

McCawley, James D. (1982) The syntax and semantics of English relative clauses. *Lingua* 53: 99–149.

McCawley, James D. (1988) Review of Chomsky (1986a). *Language* 64: 355–66.

McClelland, James L., and Karalyn Patterson (2002) Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences* 6: 465–72.

McCloskey, James (1979) *Transformational Syntax and Model Theoretic Semantics: A Case-Study in Modern Irish*. Dordrecht: Reidel.

McCloskey, James (1993) Constraints on syntactic processes. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*, vol. 1, pp. 496–506. Berlin: Mouton de Gruyter.

McCloskey, James (1996) On the scope of verb movement in Irish. *Natural Language and Linguistic Theory* 14: 47–104.

McCloskey, James, and Kenneth Hale (1984) On the syntax of person–number inflection in Modern Irish. *Natural Language and Linguistic Theory* 1: 487–533.

McCrone, John (1999) States of mind. *New Scientist* 2178: 30–33.

McEnery A., R. Xiao, and Y. Tono (2006) *Corpus-Based Language Studies*. London: Routledge.

McGilvray, James (2009) Introduction to the 3rd edn of Noam Chomsky, *Cartesian Linguistics*, pp. 1–52. Cambridge: Cambridge University Press.

Medeiros, David (2008) Optimal packing in phrase structure. *Biolinguistics* 2: 152–95.

MEDIN, DOUGLAS L., and MARGUERITE M. SCHAFFER (1978) Context theory of classification learning. *Psychological Review* 85: 207–38.

MEGERDOOMIAN, KARINE (n.d.) http://crl.nmsu.edu/Research/Projects/shiraz/ling/morph.html

MEIER, RICHARD (1980) Icons and morphemes: models of the acquisition of verb agreement in ASL. *Papers and Reports on Child Language Development* 20: 92–9.

MEILLET, ANTOINE (1912) L'Évolution des formes grammaticales. *Scientia (Rivista di scienzia)* 12, 26.6: 384–400.

MEINER, JOHANN WERNER (1971[1781]) Versuch einer an der menschlichen Sprache abgebildeten Vernunftlehre oder Philosophische und allgemeine Sprachlehre. Stuttgart: Frommann.

MEL ČUK, IGOR (1988) *Dependency Syntax: Theory and Practice*. Albany, NY: State University of New York Press.

MEL'ČUK, IGOR (1997) *Vers une Linguistique Sens-Texte*. Paris: Collège de France, Chaire Internationale.

MEL'ČUK, IGOR (2003) Levels of dependency in linguistic description: concepts and p(roblems. In Ágel et al. (2003: 188–229).

MELTZOFF, ANDREW N. (1999) Origins of theory of mind, cognition and communication. *Journal of Communication Disorders* 32: 251–69.

MELTZOFF, ANDREW N., and WOLFGANG PRINZ (eds.) (2002) *The Imitative Mind: Development, Evolution and Brain Bases*. Cambridge: Cambridge University Press.

MERCHANT, NAZARRE, and BRUCE TESAR (forthcoming) Learning underlying forms by searching restricted lexical subspaces. *Proceedings of CLS 41*. ROA- 811.

MERIN, A. (1999) Information, relevance, and social decision making: some principles and results of decision-theoretic semantics. In L. S. Moss, J. Ginzburg, and M. de Rijke (eds.), *Logic, Language, and Computation*, vol. 2, pp. 179–221. Stanford: CSLI Publications.

MESULAM, M.-MARSEL (2000) *Principles of Behavioral and Cognitive Neurology*. New York: Oxford University Press.

MEYER, C. F. (1992) *Apposition in Contemporary English*. Cambridge: Cambridge University Press.

MEYER, C. F. (2002) *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.

MICELI, GABRIELE, M. CATERINA SILVERI, GIAMPIERO VILLA, and ALFONSO CARAMAZZA (1984) On the basis for the agrammatic's difficulty in producing main verbs. *Cortex* 20: 207–20.

MICHAELIS, LAURA A. (2004) Type shifting in Construction Grammar: an integrated approach to aspectual coercion. *Cognitive Linguistics* 15: 1–67.

MICHAELIS, LAURA A., and KNUD LAMBRECHT (1996) Toward a construction-based model of language function: the case of nominal extraposition. *Language* 72: 215–47.

MICHAELIS, LAURA A., and JOSEF RUPPENHOFER (2001) *Beyond Alternations: A Construction-Based Approach to the Applicative Pattern in German*. Stanford, CA: CSLI Publications.

MIELKE, JEFF (2008) *The Emergence of Distinctive Features*. Oxford: Oxford University Press.

MILLER, JOANNE (1994) On the internal structure of phonetic categories: a progress report. *Cognition* 50: 271–85.

MILNER, A. DAVID, and MELVYN A. GOODALE (1995) *The Visual Brain in Action*. Oxford: Oxford University Press.

Mindt, D. (1995) *An Empirical Grammar of the English Verb: Modal Verbs*. Berlin: Cornelsen.
Minsky, Marvin (1975) A framework for representing knowledge. In Patrick Winston (ed.), *The Psychology of Computer Vision*, pp. 211–77. New York: McGraw-Hill.
Minsky, Marvin (1988) *The Society of Mind*. New York: Simon & Schuster.
Mioto, Carlos (1999) A periferia esquerda no português brasileiro. MS, Universidade Federal de Santa Catarina, University of Siena.
Mithun, Marianne (1987) Is basic word order universal? In Russell S. Tomlin (ed.), *Coherence and Grounding in Discourse*, pp. 281–328. Amsterdam: Benjamins.
Mithun, Marianne (1991) Active/agentive case marking and its motivations. *Language* 67: 510–46.
Mithun, Marianne (2000) Incorporation. In Geert Booij, Joachim Mugdan, and Christian Lehmann (eds.), *Morphologie/Morphology: Ein internationales Handbuch zur Flexion und Wortbildung/An International Handbook on Inflection and Word Formation*, vol. 1, pp. 916–28. Berlin: de Gruyter.
Miura, Irene T. (1987) Mathematics achievement as a function of language. *Journal of Educational Psychology* 79: 79–82.
Miura, Irene T., Chungsoon C. Kim, Chih-mei Chang, and Yukari Okamoto (1988) Effects of language characteristics on children's cognitive representation of number: cross-national comparisons. *Child Development* 59: 1445–50.
Miura, Irene T., and Yukari Okamoto (1989) Comparisons of U.S. and Japanese first graders cognitive representation of number and understanding of place value. *Journal of Educational Psychology* 81: 109–14.
Miura, Irene T., Yukari Okamoto, Chungsoon C. Kim, Chih-Mei Chang, Marcia Steere, and Michel Fayol (1994) Comparisons of children's cognitive representation of number: China, France, Japan, Korea, Sweden, and the United States. *International Journal of Behavioral Development* 17: 401–11.
Miura, Irene T., Yukari Okamoto, Chungsoon C. Kim, Marcia Steere, and Michel Fayol (1993) First graders' cognitive representation of number and understanding of place value: cross-national comparisons: France, Japan, Korea, Sweden, and the United States. *Journal of Educational Psychology* 85, 1.
Miura, Irene T., Yukari Okamoto, Vesna Vlahovic-Stetic, Chungsoon C. Kim, and Jong Hye Han (1999) Language supports for children's understanding of numerical fractions: cross-national comparisons. *Journal of Experimental Child Psychology* 74: 356–65.
Mohanan, K. P. (1982) Grammatical relations and clause structure in Malayalam. In Bresnan (1982a: 504–89).
Mok, Eva, and John Bryant (2006) A best-fit approach to productive omission of arguments. *BLS* 32.
Mönnink, I., N. Brom, and N. Oostdijk (2003) Using the MF/MD method for automatic text classification. In S. Granger and S. Petch-Tyson (eds.), *Extending the Scope of Corpus-Based Research: New Applications, New Challenges*, pp. 15–26. Amsterdam: Rodopi.
Monod, Jacques (1971) *Chance and Necessity*. New York: Knopf.
Moortgat, Michael (1997) Categorial type logics. In Johan van Benthem and Alice ter Meulen (eds.), *Handbook of Logic and Language*, pp. 93–177. Amsterdam/Cambridge, MA: Elsevier Science/MIT Press.
Moravcsik, Edith A. (1978) On the case marking of objects. In Joseph H. Greenberg (ed.), *Universals of Human Language*, pp. 249–89. Stanford, CA: Stanford University Press.

Moravcsik, Edith A. (2009) Explaining language universals. In Jae Jung Song (ed.), *Oxford Handbook of Typology.* Oxford: Oxford University Press.

Morén, Bruce (2007) The division of labor between segment-internal structure and violable constraints. In Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.), *Freedom of Analysis?* Berlin: Mouton de Gruyter.

Moreton, Elliott (2003) Non-computable functions in Optimality Theory. In John J. McCarthy (ed.), *Optimality Theory: A Reader*. Oxford: Blackwell.

Morford, Jill P. (2002) Why does exposure to language matter? In T. Givón and Bertrande Malle (eds.), *Evolution of Language out of Pre-language*, pp. 329–41. Amsterdam: Benjamins,

Morford, Jill P. (2003) Grammatical development in adolescent first language learners. *Linguistics* 41: 681–721.

Morimoto, Yukiko (2000) Discourse configurationality in Bantu morphosyntax. PhD thesis, Stanford University.

Morrill, Glyn (1994) *Type Logical Grammar: Categorial Logic of Signs*. Dordrecht: Kluwer Academic.

Morrill, Glyn (2000) Incremental processing and acceptability. *Computational Linguistics* 26: 319–38.

Morrill, Glyn, and Mario Fadda (2008) Proof nets for basic discontinuous Lambek calculus. *Logic and Computation* 18: 239–56.

Morris, Desmond, Peter Collett, Peter Marsh, and Marie O'Shaughnessy (1979) *Gestures: Their Origin and Distribution*. New York: Stein & Day.

Mosel, Ulrike, and Even Hovdhaugen (1992) *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Mufwene, Salikoko S. (2001) *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Mycock, Louise Jane (2006) The typology of constituent questions: a lexical-functional grammar analysis of 'wh'-questions. PhD thesis, University of Manchester.

Myers, Scott (1994) OCP effects in Optimality Theory. MS. Available on the Rutgers Optimality Archive, ROA 6, http://roa.rutgers.edu

Myhill, John (1995) Change and continuity in the functions of the American English modals. *Linguistics* 33: 157–211.

Myhill, John (1997) *Should* and *ought*: the rise of individually oriented modality in American English. *English Language and Linguistics* 1: 3–23.

Naess, Åshild (2007) *Prototypical Transitivity*. Amsterdam: Benjamins.

Naigles, Letitia R., and Paula Terrazas (1998) Motion-verb generalizations in English and Spanish. *Psychological Science* 9: 363–9.

Narasimhan, B. (1998) A lexical semantic explanation for 'quirky' case marking in Hindi. *Studia Linguistica* 52: 48–76.

Narayanan, Srini (1999) Moving right along: a computational model of metaphoric reasoning about events. *Proceedings of the National Conference on Artificial Intelligence* (AAAI 99, Orlando, FL), pp. 121–8.

Narayanan, Srini, and Dan Jurafsky (2001) A Bayesian model predicts parse preferences and reading times in sentence processing. In T. G. Dietterich, S. Becker, and Z. Ghahramani (eds.), *Advances in Neural Information Processing Systems*, 14, pp. 59–65. Cambridge, MA: MIT Press.

Narrog, Heiko (2004) From transitive to causative in Japanese: morphologization through exaptation. *Diachronica* 21: 351–92.

Narrog, Heiko (2007) Exaptation, grammaticalization, and reanalysis. *California Linguistic Notes* 23: 1–27.

Nesi, Hilary, and Helen Basturkmen (2006) Lexical bundles and discourse signaling in academic lectures. *International Journal of Corpus Linguistics* 11: 283–304.

Nespor, Marina, and Wendy Sandler (1999) Prosody in Israeli Sign Language. *Language and Speech* 42: 143–76.

Nespor, Marina, and Irene Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris.

Nettle, Daniel (1999) Functionalism and its difficulties in biology and linguistics. In Darnell et al. (1999: 445–67).

Neumann, Günter, and Dan Flickinger (2002) HPSG-DOP: data-oriented parsing with HPSG. *Proceedings 9th International Conference on HPSG* (HPSG-2002, Seoul).

Newmeyer, Frederick J. (1998) *Language Form and Language Function*. Cambridge, MA: MIT Press.

Newmeyer, Frederick J. (2003) Grammar is grammar and usage is usage. *Language* 79: 682–707.

Newmeyer, Frederick J. (2005) *Possible and Probable Languages*. Oxford: Oxford University Press.

Newmeyer, Frederick J. (2006) On Gahl and Garnsey on grammar and usage. *Language* 82: 399–404.

Newport, Elissa (1991) Contrasting concepts of the critical period for language. In Susan Carey and Rochel Gelman (eds.), *The Epigenesis of Mind: Essays on Biology and Cognition*, pp. 111–30. Hillsdale, NJ: Erlbaum.

Nichols, Johanna (1986) Head-marking and dependent-marking grammar. *Language* 62: 56–119.

Nicol, Janet, and David Swinney (1989) The role of structure in Coreference Assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18: 5–19.

Nicolis, Marco (2008) Reinforcers in Italian: DP-internal and left peripheral. *Proceedings of CLS* 41.

Nicolle, Steve (1997) Relevance and the effect of modal auxiliaries in logical reasoning tasks. In P. Weingartner, G. Schurz, and G. Dorn (eds.), *The Role of Pragmatics in Contemporary Philosophy: Contributions of the Austrian Ludwig Wittgenstein Society*, vol. 2, pp. 721–7. Kirchberg an Wechsel: Austrian Ludwig Wittgenstein Society.

Nicolle, Steve, and Billy Clark (1999) Experimental pragmatics and what is said: a response to Gibbs and Moise. *Cognition* 69: 337–54.

Nilsen, Øystein (2003) Eliminating positions: syntax and semantics of sentential modification. PhD dissertation, Utrecht University.

Ninio, Anat (2006) *Language And The Learning Curve: A New Theory Of Syntactic Development.* Oxford: Oxford University Press.

Nizegorodcew, Anna (2007) *Input for Instructed L2 Learners: The Relevance of Relevance.* London: Multilingual Matters.

Nolan, B. (ed.) (2004a) *RRG 2004 Book of Proceedings*. Available on RRG website: http://linguistics.buffalo.edu/research/rrg.html

Nolan, B. (2004b) First steps toward a computational RRG. In Nolan (2004a: 196–223).

Norde, Muriel (2001) Deflexion as a counterdirectional factor in grammatical change. *Language Sciences* 23: 231–64.

NORDLINGER, RACHEL (1998) *Constructive Case: Evidence from Australian Languages*. Stanford, CA: CSLI Publications.

NOSOFSKY, ROBERT M. (1988) Similarity, frequency and category representation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14: 54–65.

NOTMAN, LESLIE A., PAUL T. SOWDEN, and EMRE ÖZGEN (2005) The nature of learned categorical perception effects: a psychophysical approach. *Cognition* 95: B1–B14.

NUNBERG, GEOFFREY, IVAN A. SAG, and THOMAS WASOW (1994) Idioms. *Language* 70: 491–538.

NUNBERG, GEOFFREY D., and ANNIE ZAENEN (1992) Systematic polysemy in lexicology and lexicography. *Proceedings of Euralex II* 2: 387–98.

NUYTS, J. (1993) On determining the functions of language. *Semiotica* 94: 201–32.

O'CONNOR, ROBERT (2006) Information structure in Lexical-Functional Grammar: the discourse–prosody correspondence. PhD thesis, University of Manchester.

ODDEN, DAVID (1988) Anti antigemination and the OCP. *Linguistic Inquiry* 19: 451–75.

O'DONOVAN, RUTH, MICHAEL BURKE, AOIFE CAHILL, JOSEF VAN GENABITH, and ANDY WAY (2005) Large-scale induction and evaluation of lexical resources from the Penn-II and Penn-III treebanks. *Computational Linguistics* 31: 328–65.

OEZGEN, EMRE, and IAN R. L. DAVIES (1998) Turkish color terms: tests of Berlin and Kay's theory of color universals and linguistic relativity. *Linguistics* 36: 919–56.

O'GRADY, WILLIAM (1997) *Syntactic Development*. Chicago: University of Chicago Press.

O'GRADY, WILLIAM (2005) *Syntactic Carpentry: An Emergentist Approach to Syntax*. Mahwah, NJ: Erlbaum.

O'GRADY, WILLIAM (2008a) Does emergentism have a chance? In H. Chan, H. Jacob, and E. Kapia (eds.), *Proceedings of 32nd Annual Boston University Conference on Language Development*, pp. 16–35. Sommerville, MA: Cascadilla Press.

O'GRADY, WILLIAM (2008b) Innateness, Universal Grammar, and emergentism. *Lingua* 118: 620–31.

O'GRADY, WILLIAM (2008c) Language without grammar. In P. Robinson and N. Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, pp. 139–67. New York: Routledge.

O'GRADY, WILLIAM (2008d) The Emergentist Program. *Lingua* 118: 447–64.

OH, K. J. (2003) Language, cognition, and development: motion events in English and Korean. PhD dissertation, University of California, Berkeley.

O'HALLORAN, KAY L. (ed.) (2006) *Multimodal Discourse Analysis: Systemic Functional Perspectives*. London: Continuum.

OHARA, KYOKO, SEIKO FUJII, SHUN ISHIZAKI, TOSHIO OHORI, HIROAKI SATO, and RYOKO SUZUKI (2004) The Japanese FrameNet project: an introduction. In Charles J. Fillmore, Manfred Pinkal, Collin F. Baker, and Katrin Erk (eds.), *Proceedings of LREC 2004, Workshop on Building Lexical Resources from Semantically Annotated Corpora* (BLRSAC, Lisbon), pp. 9–12.

OHORI, T. (1992) Diachrony in clause linkage and related issues. PhD dissertation, UC Berkeley. Available on RRG website: http://linguistics.buffalo.edu/research/rrg.html

OHORI, T. (2001) Some thoughts on a new systematization of interclausal semantic relations. Paper presented at 2001 Role and Reference Grammar Conference (University of California, Santa Barbara).

Ohori, T. (2005) More thoughts on the semantic representation in RRG: event types and the semantics of clause linkage. Paper presented at the 2005 International RRG Conference, Taiwan. Available on RRG website: http://linguistics.buffalo.edu/research/rrg.html

Olson, M. (1981) Barai clause junctures: toward a functional theory of interclausal relations. PhD dissertation, Australian National University.

Ono, Tsuyoshi, Ryoko Suzuki, and Sandra A. Thompson (2000) The pragmatic nature of the so-called subject marker *ga* in Japanese: evidence from conversation. *Discourse Studies* 2: 55–84.

Orgun, C. Orhan, and Ronald Sprouse (1999) From MParse to control: deriving ungrammaticality. *Phonology* 16: 191–220.

Osawa, Fuyo (2003) Syntactic parallels between ontogeny and phylogeny. *Lingua* 113: 3–47.

Osterhout, Lee, and Linda Mobley (1995) Event-related potentials elicited by failure to agree. *Journal of Memory and Language* 34: 739–73.

Otero, Carlos P. (1990) Prolegomena to a future evolution metaphysics: the emergence of homo loquens and physical law.*Behavioral and Brain Sciences* 12: 747–50.

Packard, Jerome (2000) *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.

Pagliuca, William (ed.) (1994) *Perspectives on Grammaticalization*. Amsterdam: Benjamins.

Pagliuca, William, and Richard Mowrey (1987) Articulatory evolution. In Anna G. Ramat, Onofrio Carruba, and Giuliano Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*, pp. 459–72. Amsterdam: Benjamins.

Panther, Klaus-Uwe, and Linda L. Thornburg (2003) Introduction: on the nature of conceptual metonymy. In Klaus-Uwe Panther and Linda L. Thornburg (eds.), *Metonymy and Pragmatic Inferencing*, pp. 1–20. Amsterdam: Benjamins.

Papafragou, Anna, Chris Massey, and Lila Gleitman (2001) Shake, rattle, 'n' roll: the representation of motion in language and cognition. *Cognition* 84: 189–219.

Paradis, Michel (2004) *A Neurolinguistic Theory of Bilingualism*. Amsterdam: Benjamins.

Parker, A. (2006) Evolution as a constraint on theories of syntax: the case against minimalism. Doctoral dissertation, University of Edinburgh.

Parker, Steve (2008) Sound level protrusions as physical correlates of sonority. *Journal of Phonetics* 1: 55–90.

Parodi, Giovanni (2007) Variation across registers in Spanish. In G. Parodi (ed.), *Working with Spanish Corpora*, pp. 11–53. London: Continuum.

Parsons, Terence (1990) *Events in the Semantics of English: A Study in Subatomic Semantics*. Cambridge, MA: MIT Press.

Partee, Barbara H., and Vladimir Borschev (2007) Existential sentences, BE and the genitive of negation in Russian. In K. von Heisenger and I. Comorovski (eds.), *Existence: Semantics and Syntax*, pp. 147–90. Berlin: Springer.

Partee, Barbara H., Alice ter Meulen, and Robert E. Wall (1993) *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.

Partington, A., and J. Morley (2004) From frequency to ideology: investigating word and cluster/bundle frequency in political debate. In B. Lewandowska-Tomaszczyk (ed.), *Practical Applications in Language and Computers* (PALC 2003), pp. 179–92. Frankfurt a. Main: Lang.

PASIERBSKY, FRITZ (1981) Sprachtypologische Aspekte der Valenztheorie unter besonderer Berücksichtigung des Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 34: 160–77.

PAUL, HERMANN (1888) *Principles of the History of Language*. London: Sonnenschein & Co.

PAWLEY, ANDREW (2006) Where have all the verbs gone? Remarks on the organisation of languages with small, closed verb classes. Paper presented at the 11th Biennenial Rice University Linguistics Symposium, 16–18 March. (http://www.ruf.rice.edu/∼lingsymp/Pawley_paper.pdf)

PAWLEY, ANDREW, and FRANCES HODGETTS SYDER (1983) Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In Jack C. Richards and Richard W. Schmidt (eds.), *Language and Communication*, pp. 191–226. New York: Longman.

PAYNE, THOMAS E. (1997) *Describing Morphosyntax: a Guide for Field Linguists*. Cambridge: Cambridge University Press.

PEDERSON, ERIC (1995) Language as context, language as means: spatial cognition and habitual language use. *Cognitive Linguistics* 6: 33–62.

PEDERSON, ERIC (2007) Cognitive linguistics and linguistic relativity. In D. Geeraerts and H. Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, pp. 1012–44. Oxford: Oxford University Press.

PEDERSON, ERIC, EVE DANZIGER, STEPHEN LEVINSON, SOTARO KITA, GUNTER SENFT, and DAVID WILKINS (1998) Semantic typology and spatial conceptualization. *Language* 74: 557–89.

PEETERS, BERT (2000) 'S'engager' vs. 'to show restraint: linguistic and cultural relativity in discourse management. In Susanne Niemeier and René Dirven (eds.), *Evidence for Linguistic Relativity*, pp. 193–222. Amsterdam: Benjamins.

PEETERS, BERT (ed.) (2006) *Semantic Primes and Universal Grammar: Empirical Evidence from the Romance Languages*. Amsterdam: Benjamins.

PELCZAR, MICHAEL W. (2004) The indispensability of *Farbung*. *Synthese* 138: 49–77.

PELCZAR, MICHAEL W. (2007) Forms and objects of thought. *Linguistics and Philosophy* 30: 97–122.

PELLETIER, JEFF, and RENEE ELIO (2005) The case for psychologism in default and inheritance reasoning. *Synthese* 146: 7–35.

PENKE, MARTINA (2001) Controversies about CP: a comparison of language acquisition and language impairments in Broca's Aphasia. *Brain and Language* 77: 351–63.

PENTUS, MATI (1993) *Lambek calculus is L-complete*. ILLC Report, University of Amsterdam. Shortened version published 1994 as 'Language completeness of the Lambek calculus', *Proceedings of the Ninth Annual IEEE Symposium on Logic in Computer Science* (Paris), pp. 487–96.

PEPPERBERG, IRENE M. (1999) *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Cambridge, MA: Harvard University Press.

PERELTSVAIG, ASYA (2007) On the universality of DP: a view from Russian. *Studia Linguistica* 61: 59–94.

PÉREZ, AVELINE (1990) Time in motion: grammaticalisation of the *be going to* construction in English. *La Trobe University Working Papers in Linguistics* 3:49–64.

PERFORS, AMY, JOSHUA TENENBAUM, and TERRY REGIER (2006) Poverty of the stimulus? A rational approach. *Proceedings 28th Annual Conference of the Cognitive Science Society* (Vancouver).

Perlmutter, David M. (1980) Relational grammar. In Edith A. Moravcsik and Jessica R. Wirth (eds.), *Current Approaches to Syntax.* San Diego: Academic Press, pp. 195–229.

Perlmutter, David M. (ed.) (1983) *Studies in Relational Grammar 1.* Chicago: University of Chicago Press.

Perlmutter, David M. (1989) A moraic theory of American Sign Language syllable structure. MS.

Perlmutter, David M., and Paul M. Postal (1977) Toward a universal characterization of passivization. *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*. Reprinted in Perlmutter (1983: 3–29).

Perlmutter, David M., and Paul M. Postal (1983) Toward a universal characterization of passivization. In Perlmutter (1983: 1–29).

Perner, Josef, Uta Frith, Alan M. Leslie, and Sue Leekam (1989) Exploration of the autistic child's theory of mind: knowledge, belief and communication. *Child Development* 60: 689–700.

Pesetsky, David (1997) Optimality theory and syntax: movement and pronunciation. In D. Archangeli and D. T. Langendoen (eds.), *Optimality Theory: An Overview*, pp. 134–70. Oxford: Blackwell.

Peters, Ann (1983) *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

Petri, Herbert L., and Mortimer Mishkin (1994) Behaviorism, cognitivism and the neuropsychology of memory. *American Scientist* 82: 30–37.

Pfau, Roland, and Markus Steinbach (2005) Grammaticalization of auxiliaries in sign languages. TS, University of Amsterdam and University of Mainz.

Pfau, Roland, and Markus Steinbach (2006) Modality-independent and modality-specific aspects of grammaticalization in sign languages. *Linguistics in Potsdam* 24: 5–98.

Phillips, Colin (2006) The real-time status of island phenomena. *Language* 82: 795–823.

Pianesi, Fabio, and Achille C. Varzi (2000) Events and event talk: an introduction. In J. Higginbotham, F. Pianesi, and A. C. Varzi (eds.), *Speaking of Events*, pp. 3–47. Oxford: Oxford University Press.

Piantadosi, Steven, Noah Goodman, Benjamin Ellis, and Joshua Tenenbaum (2008) A Bayesian model of the acquisition of compositional semantics. *Proceedings CogSci 2008*.

Piattelli-Palmarini, Massimo (1980) *Language and Learning*. Cambridge, MA: Harvard University Press.

Piattelli-Palmarini, Massimo (1989) Evolution, selection and cognition: from learning to parameter setting in biology and in the study of language.*Cognition* 31: 1–44.

Piattelli-Palmarini, Massimo (2006) The return of the laws of form. C. P. Manzù (ed.), *Life on the Edge* (*La Vita in Bilico*), vol. 2 (Workshop 1) (Rome), pp. 45–57.

Pickering, Martin, and Matthew Traxler (2001) Strategies for processing unbounded dependencies: lexical information and verb-argument assignment.*Journal of Experimental Psychology* 27: 1401–10.

Pierrehumbert, Janet (2001) Exemplar dynamics: word frequency, lenition and contrast. In Bybee and Hopper (2001: 137–57).

Pierrehumbert, Janet (2002) Word-specific phonetics. In Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory Phonology 7*, pp. 101–39. Berlin: Mouton de Gruyter.

PIETROSKI, PAUL M. (2003) Small verbs, complex events: analyticity without synonymy. In Louise M. Antony and Norbert Hornstein (eds.), *Chomsky and his Critics*, pp. 179–214. Oxford: Blackwell.

PIETROSKI, PAUL M. (2006) Interpreting concatenation and concatenates. *Philosophical Issues* 16: 221–45.

PIETROSKI, PAUL M. (forthcoming) *Semantics Without Truth-Values*. Oxford: Oxford University Press.

PIGLIUCCI, MASSIMO (2007a) Postgenomic musings. *Science* 317: 1172–3.

PIGLIUCCI, MASSIMO (2007b) Do we need an extended evolutionary synthesis? *Evolution* 61.12: 2743–9.

PILKINGTON, ADRIAN (2000) *Poetic Effects: A Relevance Theory Perspective*. Amsterdam: Benjamins.

PINKER, STEVEN (1984) *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

PINKER, STEVEN (1989) *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

PINKER, STEVEN (1991) Rules of language. *Science* 253: 530–35.

PINKER, STEVEN (1997) *How the Mind Works*. New York: Norton.

PINKER, STEVEN (1999) *Words and Rules*. New York: Basic Books.

PINKER, STEVEN, and PAUL BLOOM (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13: 707–26.

PINKER, STEVEN, and RAY JACKENDOFF (2005). The faculty of language: what's special about it? *Cognition* 95: 201–36.

PITTS, ALYSON (2005) Assessing the evidence for intuitions about *what is said*. MS, University of Cambridge.

PIZZUTO, ELENA (2004[1987]) Aspetti morfosintattici. In Virginia Volterra (ed.), *La lingua italiana dei segni: la comunicazione visivo-gestuale dei sordi*, pp. 179–209. Bologna: Il Mulino.

PLAG, INGO (1996) Selectional restrictions in English suffixation revisited. *Linguistics* 34: 769–98.

PLANK, FRANS (1984) Verbs and objects in semantic agreement: minor differences between languages that might suggest a major one. *Journal of Semantics* 3: 305–60.

POEPPEL, DAVID, and DAVID EMBICK (2005) The relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, pp. 103–18. Mahwah, NJ: Erlbaum.

POLETTO, CECILIA (2000) *The Higher Functional Field: Evidence from Northern Italian Dialects*. New York: Oxford University Press.

POLLARD, CARL, and IVAN SAG (1992) Anaphora in English and the scope of Binding Theory. *Linguistic Inquiry* 23: 261–303.

POLLARD, CARL, and IVAN SAG (1994) *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

POLLOCK, JEAN-YVES (1989) Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365–424.

POPLACK, SHANA (2001) Variability, frequency and productivity in the irrealis domain of French. In Bybee and Hopper (2001: 405–28).

POPLACK, SHANA, and SALI TAGLIAMONTE (1999) The grammaticization of *going to* in (African American) English. *Language Variation and Change* 11: 315–42.

Poplack, Shana, and Sali Tagliamonte (2001) *African American English in the Diaspora*. Oxford: Blackwell.

Posner, Michael I., and Jin Fan (2004) Attention as an organ system. In J. Pomerantz (ed.), *Neurobiology of Perception and Communication: From Synapses to Society* (the IVth De Lange Conference). Cambridge: Cambridge University Press.

Posner, Michael I., and Steven W. Keele (1968) On the genesis of abstract ideas. *Journal of Experimental Psychology* 77: 353–63.

Posner, Michael I., and Antonella Pavese (1997) Anatomy of word and sentence meanings. Paper presented at the Colloquium on Neuroimaging of Human Brain Functions, organized by M. Posner and M. Raichle, National Academy of Science, USA (Irvine, CA, May).

Postal, Paul M. (1969) Anaphoric islands. In Robert I. Binnick, Alice Davison, Georgia Green, and Jerry L. Morgan (eds.), *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, pp. 205–35. Chicago: University of Chicago.

Postal, Paul (1972) The best theory. In P. S. Peters (ed.), *Goals of Linguistic Theory*, pp. 131–70. Englewood Cliffs, NJ: Prentice Hall.

Pottier, Bernard (1992) *Sémantique générale.* Paris: PUF.

Povinelli, Daniel J., and Todd M. Preuss (1995) Theory of mind: evolutionary history of a cognitive specialization. *Trends in Neuroscience* 18: 418–24.

Prawitz, Dag (1965) *Natural Deduction*. Stockholm: Almqvist & Wiksell.

Predelli, Stefano (2005a) *Contexts: Meaning, Truth, and the Use of Language*. Oxford: Clarendon Press.

Predelli, Stefano (2005b) Painted leaves, context, and semantic analysis. *Linguistics and Philosophy* 28: 351–74.

Premack, David (1971) Language in chimpanzees. *Science* 172: 808–22.

Premack, David, and Guy Woodruff (1978) Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1: 515–26.

Preyer, Gerhard, and Georg Peter (eds.) (2005) *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Oxford: Clarendon Press.

Priestley, Carol (2008) The semantics of inalienable possession in Koromu (PNG). In Cliff Goddard (ed.), *Cross-Linguistic Semantics*, pp. 277–300. Amsterdam: Benjamins.

Primus, Beatrice (1999) *Case and Thematic Roles: Ergative, Accusative and Active*. Tübingen: Niemeyer.

Prince, Alan (1985) Improving tree theory. In M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Jeder (eds.), *Proceedings of Berkeley Linguistics Society 11*, pp. 471–90. Berkeley, CA: Berkeley Linguistic Society.

Prince, Alan (1998) Paninian relations. Handout, University of Marburg. Available online at http://ling.rutgers.edu/gamma/talks/panrel-mbg.pdf.

Prince, Alan, and Paul Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report no. 2, RuCCS, Rutgers University, New Brunswick, NJ.

Prince, Alan, and Paul Smolensky (1997) Optimality: from neural networks to universal grammar. *Science* 275: 1604–10.

Prince, Alan, and Paul Smolensky (2002) Optimality Theory in phonology. In *Oxford International Encyclopedia of Linguistics*. Oxford: Oxford University Press.

Prince, Alan, and Paul Smolensky (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

Prince, Alan, and Bruce Tesar (2004) Learning phonotactic distributions. In René Kager, Joe Pater, and Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, pp. 245–91. Cambridge: Cambridge University Press.

Prince, E. F. (1978) A comparison of *wh*-clefts and *it*-clefts in discourse. *Language* 54: 883–906.

Progovac, Ljiljana (1998) Determiner phrase in a language without determiners. *Journal of Linguistics* 34: 165–79.

Pullum, Geoffrey (1982) Syncategorematicity and English infinitival *to*. *Glossa* 16: 181–215.

Pullum, Geoffrey, and Arnold Zwicky (1991) Condition duplication, paradigm homonymy and transconstructional constraints. *Berkeley Linguistics Society* 17: 252–66.

Pulvermüller, Friedemann (2000) Syntactic circuits: how does the brain create serial order in sentences? *Brain and Language* 71: 194–9.

Purvis, Tristan M. (2008) A linguistic and discursive analysis of register variation in Dagbani. PhD dissertation, Indiana University.

Pustejovsky, James (1995) *The Generative Lexicon*. Cambridge, MA: MIT Press.

Putnam, Hilary (1962) The analytic and the synthetic. *Minnesota Studies in the Philosophy of Sciences* 3: 358–97.

Pütz, Martin, Susanne Niemeier, and René Dirven (eds.) (2001a) *Applied Cognitive Linguistics I: Theory and Language Acquisition*. Berlin: Mouton de Gruyter.

Pütz, Martin, Susanne Niemeier, and René Dirven (eds.) (2001b) *Applied Cognitive Linguistics II: Language Pedagogy*. Berlin: Mouton de Gruyter.

Pütz, Martin, and Marjolijn Verspoor (2000) *Explorations in Linguistic Relativity*. Amsterdam: Benjamins.

Quaglio, Paulo (2009) *Television Dialogue: The Sitcom 'Friends' versus Natural Conversation*. Amsterdam: Benjamins.

Quillian, Ross, and Collins, Allan (1969) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8: 240–47.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*. Harlow: Longman.

Radford, Andrew (1990) *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Oxford: Blackwell.

Rainer, Franz (1996) Inflection inside derivation: evidence from Spanish and Portuguese. *Yearbook of Morphology* 1995: 83–91.

Rajagopalan, Kanavillil (2005) Linguistic pragmatics: a discipline bedeviled by its own history? Implications for intercultural studies. *Intercultural Pragmatics* 2: 93–7.

Ramat, Anna Giacalone, and Paul J. Hopper (eds.) (1998) *The Limits of Grammaticalization*. Amsterdam: Benjamins.

Ramat, Paolo (1992) Thoughts on degrammaticalization. *Linguistics* 30: 549–60.

Ramchand, Gillian (2008) *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge: Cambridge University Press.

Rappaport Hovav, Malka, and Beth Levin (1998) Building verb meanings. In Miriam Butt and Wilhelm Geuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, pp. 97–134. Stanford, CA: CSLI Publications.

Reali, Florencia, and Morten Christiansen (2005) Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cognitive Science* 29: 1007–28.

Recanati, François (1989) The pragmatics of what is said. *Mind and Language* 4: 295–329. Reprinted in S. Davis (ed.), *Pragmatics: A Reader*, pp. 97–21. Oxford: Oxford University Press, 1991.

Recanati, François (1993) *Direct Reference: From Language to Thought*. Oxford. Blackwell.

Recanati, François (2002) Unarticulated constituents. *Linguistics and Philosophy* 25: 299–345.

Recanati, François (2004a) *Literal Meaning*. Cambridge: Cambridge University Press.

Recanati, François (2004b) Pragmatics and semantics. In Horn and Ward (2004: 442–62).

Recanati, François (2005) Literalism and contextualism: some varieties. In G. Preyer and G. Peter (eds.), *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, pp. 171–96. Oxford: Clarendon Press.

Recanati, François (2007) Reply to Carston 2007. In M. J. Frápolli (ed.), *Saying, Meaning and Referring: Essays on François Recanati's Philosophy of Language*, pp. 49–54. Basingstoke: Palgrave Macmillan.

Reddy, Michael (1979) The conduit metaphor: a case of frame conflict in our language about language. In Andrew Ortony (ed.), *Metaphor and Thought*, pp. 164–201. Cambridge: Cambridge University Press.

Redington, Mark, Nick Chater, and Steven Finch (1998) Distributional information: a powerful cue for acquiring syntactic categories. *Cognitive Science* 22: 425–69.

Regier, Terry (1996) *The Human Semantic Potential: Spatial Language and Constrained Connectionism*. Cambridge, MA: MIT Press.

Reichmann, Oskar (1988) Zur Vertikalisierung des Varietätenspektrums in der jüngeren Sprachgeschichte des Deutschen. In Horst Haider Munske, Peter von Polenz, Oskar Reichmann, and Reiner Hildebrandt (eds.), *Deutscher Wortschatz. Lexikologische Studien: Festschrift für Ludwig Erich Schmitt von seinen Marburger Schülern*, pp. 151–80. Berlin: de Gruyter.

Reinhart, Tanya (1983) *Anaphora and Semantic Interpretation*. Chicago: University of Chicago Press.

Reinhart, Tanya (2005) *Interface Strategies: Reference-Set Computation*. Cambridge, MA: MIT Press.

Reinhart, Tanya (2006) *Interface Strategies*. Cambridge, MA: MIT Press.

Reinhart, Tanya, and Eric Reuland (1993) Reflexivity. *Linguistic Inquiry* 24: 657–720.

Reisberg, Daniel (2007) *Cognition: Exploring the Science of the Mind*, 3rd media edn. New York: Norton.

Reppen, Randi (2001) Register variation in student and adult speech and writing. In Conrad and Biber (2001: 187–99).

Resches, Mariela, and Miguel Perez Pereira (2007) Referential communication abilities and theory of mind development in preschool children. *Journal of Child Language* 34: 21–52.

Reuland, Eric (2006) Binding conditions: how are they derived? MS, University of Utrecht.

Rey, Jennifer M. (2001) Historical shifts in the language of women and men: gender differences in dramatic dialogue. In Conrad and Biber (2001: 138–56).

Rialland, Annie, and Stéphane Robert (2001) The intonational system of Wolof. *Linguistics* 39: 893–939.

Rice, Curt (2007) The role of Gen and Con in modeling ternary rhythm. In Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.), *Freedom of Analysis?* Berlin: Mouton de Gruyter, 235–56.

Rice, Keren (2000) *Morpheme Order and Semantic Scope: Word Formation in the Atahapaskan Verb*. Cambridge: Cambridge University Press.

Richards, Norvin (1997) What moves where when in which language? PhD dissertation, MIT.

Riezler, Stefan, and Mark Johnson (2001) Statistical models of language learning and use. *Cognitive Science* 26: 3.

Rijkhoff, Jan (2002) *The Noun Phrase*. Oxford: Oxford University Press.

Rivelis, Eugene (2007) *Kak Vozmožen Dvujazyčnyj Slovar* [How is the Bilingual Dictionary Possible?]. Stockholm: Stockholm University.

Rizzi, Luigi (1985) Two notes on the linguistic interpretation of Broca's Aphasia. In M.-L. Kean (ed.), *Agrammatism*, pp. 153–64. New York: Academic Press.

Rizzi, Luigi (1990) *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi (1997) The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of Grammar*, pp. 281–337. Amsterdam: Kluwer.

Rizzi, Luigi (2001) On the position 'Int(errogative)' in the left periphery of the clause. In G. Cinque and G. Salvi (eds.), *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*, pp. 287–96. Amsterdam: Elsevier North-Holland.

Rizzi, Luigi (ed.) (2004a) *The Structure of CP and IP: The Cartography of Syntactic Structures*, vol. 2. New York: Oxford University Press.

Rizzi, Luigi (2004b) Locality and left periphery. In A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3. New York: Oxford University Press.

Rizzi, Luigi, and Ur Shlonsky (2007) Strategies of subject extraction. In H.-M. Gärtner and U. Sauerland (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, pp. 115–16. Berlin: Mouton de Gruyter.

Rizzolatti, Giacomo, and Michael A. Arbib (1998) Language within our grasp. *Trends in Neuroscience* 21: 188–94.

Rizzolatti, Giacomo, L. Fadiga, L. Fogassi, and V. Gallese (1996) Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3: 131–41.

Rizzolatti, Giacomo, Fadiga, L., Matelli, M., Bettinardi, V., Paulescu, E., Perani, D., and Fazio, F. (1996) Localization of grasp representation in humans by PET: 1: Observation versus execution. *Experimental Brain Research* 111: 246–51.

Rizzolatti, G., L. Fogassi, and V. Gallese (2000) Cortical mechanisms subserving object grasping and action recognition: a new view on the cortical motor functions. MS.

Roberts, Kenneth (1983) Comprehension and production of word order in Stage I. *Child Development* 54: 443–9.

Robins, Robert (2001) In Defence of WP. *Transactions of the Philological Society* 99: 114–44. (Reprinted from *TPhS* 1959.)

Roca, Iggy (ed.) (1997) *Derivations and Constraints in Phonology*. Oxford: Oxford University Press.

Rohdenburg, Günter, and Britta Mondorf (eds.) (2003) *Determinants of Grammatical Variation in English*. Berlin: Mouton de Gruyter.

Roland, D., J. L. Elman, and V. S. Ferreira (2005) Why is *that*? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition* 20: 1–28.

Römer, Ute (2005) *Progressives, Patterns, Pedagogy: A Corpus-Driven Approach to English Progressive Forms, Functions, Contexts and Didactics.* Amsterdam: Benjamins.

Roorda, Dirk (1991) Resource logics: proof-theoretical investigations. PhD thesis, University of Amsterdam.

Rosch, Eleanor (1975) Cognitive representation of semantic categories. *Journal of Experimental Psychology: General* 104: 192–233.

Rosch, Eleanor (1978) Principles of categorization. In Eleanor Rosch and B.B. Lloyd (eds.), *Cognition and Categorization*, pp. 27–48. Hillsdale, NJ: Erlbaum.

Rosch, Eleanor, and Carolyn B. Mervis (1975) Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7: 573–605.

Rosenbach, Anette (2006) Descriptive genitives in English: a case study on constructional gradience. *English Language and Linguistics* 10: 77–118.

Rosenbach, Anette (2007) Emerging variation: determiner genitives and noun-modifers. *English Language and Linguistics* 11: 143–89.

Rosenbaum, Peter (1967) *The Grammar of English Predicate Complement Constructions.* Cambridge, MA: MIT Press.

Rosenbaum, Peter S. (1970) A principle governing deletion in English sentential complementation. In Roderick A. Jacobs and Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, pp. 20–29. Waltham, MA: Ginn.

Ross, Brian H., and Valerie S. Makin (1999) Prototype versus exemplar models in cognition. In Robert J. Sternberg (ed.), *The Nature of Cognition*, pp. 205–41. Cambridge, MA: MIT Press.

Ross, John R. (1967) Constraints on variables in syntax. PhD dissertation, MIT.

Ross, John R. (1969) Guess who. In Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan (eds.), *Proceedings of the Fifth Annual Meeting of CLS*, pp. 252–86. Chicago Linguistics Society.

Ross, John R. (1973) Slifting. In Maurice Gross, Morris Halle, and M.-P. Schutzenberger (eds.), *The Formal Analysis of Natural Language*, pp. 133–69. The Hague: Mouton. RRG website: http://linguistics.buffalo.edu/research/rrg.html

Rumelhart, David E. (1975) Notes on a schema for stories. In Daniel G. Bobrow and A. Collins (eds.), *Representation and Understanding: Studies in Cognitive Science*, pp. 185–210. New York: Academic Press.

Runner, Jeffrey, Rachel Sussman, and Michael Tanenhaus (2006) Processing reflexives and pronouns in picture noun phrases. *Cognitive Science* 30: 193–241.

Sadler, Louisa, and Andrew Spencer (eds.) (2004) *Projecting Morphology*. Stanford, CA: CSLI Publications.

Sadock, Jerrold (1991) *Autolexical Syntax: A Theory Of Parallel Grammatical Representations.* Chicago: University of Chicago Press.

Saffran, Jenny R., Richard N. Aslin, and Elissa L. Newport (1996) Statistical learning by 8-month-old infants. *Science* 274: 1926–8.

Saffran, Jenny R., Elizabeth K. Johnson, Richard N. Aslin, and Elissa L. Newport (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70: 27–52.

Sag, Ivan A. (2007) Sign-based Construction Grammar: an informal synopsis. MS, Stanford University.

Sag, Ivan A. (2008) English filler–gap constructions. MS, Stanford University.

Sag, Ivan A., and Carl Pollard (1991) An integrated theory of complement control. *Language* 67: 63–113.

Sag, Ivan A., and Thomas Wasow (1999) *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications.

Sag, Ivan A., Thomas Wasow, and Emily Bender (2003) *Syntax: A Formal Introduction*. Stanford, CA: CSLI Publications.

Sáiz, M. (1999) A cross-linguistic corpus-based analysis of linguistic variation. PhD dissertation, University of Manchester.

Sakai, Katsuyuki, Okihide Hikosaka, and Kae Nakamura (2004) Emergence of rhythm during motor learning. *Trends in Cognitive Sciences* 8: 547–53.

Salem, A. (1987) *Pratique des segments répétés*. Paris: Institut National de la Langue Française.

Samuels, Bridget (2009) The structure of phonological theory. Doctoral dissertation, Harvard University.

Sandler, Wendy (1986) The spreading hand autosegment of American Sign Language. *Sign Language Studies* 50: 1–28.

Sandler, Wendy (1999) Prosody in two natural language modalities. *Language and Speech* 42: 127–42.

Sankoff, Gillian. (2004) Adolescents, young adults and the critical period: two case studies from Seven Up. In Carmen Fought (ed.), *Sociolinguistic Variation: Critical Reflections*, pp. 121–39. Oxford: Oxford University Press.

Sankoff, Gillian, and Hélène Blondeau (2007) Language change across the lifespan:/r/in Montreal French. *Language* 83: 560–88.

Sankoff, Gillian, and Penelope Brown (1976) On the origins of syntax in discourse: a case study of Tok Pisin relatives. *Language* 52: 631–66.

Sato, Hiroaki (2008) New functions of FrameSQL for multilingual FrameNets. *Proceedings of LREC 08*, pp. 28–30.

Sauerland, U. (2004) Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27: 367–91.

Saul, Jennifer (2002) What is said and psychological reality: Grice's project and relevance theorists' criticisms. *Linguistics and Philosophy* 25: 347–72.

Saussure, Ferdinand de (1916) *Cours de linguistique générale*. Lausanne: Payot.

Saussure, Ferdinand de (1959) *Course in General Linguistics*. New York: Philosophical Library.

Savage, Ceri, Elena Lieven, Anna Theakston, and Michael Tomasello (2003) Testing the abstractness of children's linguistic representations: lexical and structural priming of syntactic constructions in young children. *Developmental Science* 6: 557–67.

Savage-Rumbaugh, Susan, and R. Lewin (1994) *Kanzi: The Ape at the Brink of the Human Mind*. New York: Wiley.

Savage-Rumbaugh, Susan, J. Murphy, R. A. Sevcik, K. E. Brakke, S. L. Williams, and D. M. Rumbaugh (1993) *Language Comprehension in Ape and Child*. Ann Arbor, MI: Society for Research in Child Development.

Saxton, Matthew, and John N. Towse (1998) Linguistic relativity: the case of place value in multi-digit numbers. *Journal of Experimental Child Psychology* 69: 66–79.

Scha, Remo, Rens Bod, and Khalil Sima'an (1999) A memory-based model of syntactic analysis: data-oriented parsing. *Journal of Experimental and Theoretical Artificial Intelligence* 11: 409–40.

Schachter, Paul (1978) Review of Richard Hudson, *Arguments for a Non-transformational Grammar*. *Language* 54: 348–76.

Schachter, Paul (1981) Daughter-dependency grammar. In E. Moravcsik and J. Wirth (eds.), *Current Approaches to Syntax*, pp. 267–300. New York: Academic Press.

Schachter, Paul (1985) Parts-of-speech systems. In T. Shopen (ed.), *Language Typology: A Syntactic Description*. vol. 1: *Clause Structure*, pp. 3–61. Cambridge: Cambridge University Press.

Schachter, Paul, and Fé T. Otanes (1972) *Tagalog Reference Grammar.* Berkeley: University of California Press.

Schachter, Paul, and Timothy Shopen (2007) Parts-of-speech systems. In T. Shopen (ed.), *Language Typology: A Syntactic Description*, vol. 1: *Clause Structure*, 2nd edn, pp. 1–60. Cambridge: Cambridge University Press.

Schank, Roger, and Robert Abelson (1977) *Scripts, Plans, Goals And Understanding: An Inquiry Into Human Knowledge Structures*. Hillsdale, NJ: Erlbaum.

Schiffer, Stephen (1991) Does Mentalese have a compositional semantics? In B. Loewer and G. Rey (eds.), *Meaning in Mind: Fodor and his Critics*, pp. 181–99. Oxford: Blackwell.

Schiffer, Stephen (1994) A paradox of meaning. *Noûs* 28: 279–324.

Schiffer, Stephen (2003) *The Things We Mean*. Oxford: Clarendon Press.

Schiffrin, Deborah (1981) Tense variation in narrative.*Language* 57: 45–62.

Schiffrin, Deborah (1985a) Multiple constraints on discourse options: a quantitative analysis of causal sequences. *Discourse Processes* 8: 281–303.

Schiffrin, Deborah (1985b) Conversational coherence: the role of *well*. *Language* 61: 640–67.

Schmidt, Jürgen Erich (1991) Konvergenzen zwischen neueren Grammatiktheorien und Deskriptionsgrammatiken? Zum Verhältnis von Konstituenz, Rektion (Government), Valenz und Dependenz. In Elisabeth Feldbusch, Reiner Pogarell, and Cornelia Weiß (eds.), *Neue Fragen der Linguistik: Akten des 25. Linguistischen Kolloquiums, Paderborn 1990. Vol. 1: Bestand und Entwicklung*, pp. 211–18. Tübingen: Niemeyer.

Schmitt, Cristina, and Alan Munn (1999) Against the Nominal Mapping Parameter: bare nouns in Brazilian Portuguese. *Proceedings of NELS* 29: 339–53.

Schmitt, N., S. Grandage, and S. Adolphs (2004) Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (ed.), *Formulaic Sequences*, pp. 127–52. Amsterdam: Benjamins.

Schneider, Walter, and Jason M. Chein (2003) Controlled and automatic processing: behavior, theory, and biological mechanisms. *Cognitive Science* 27: 525–59.

Schneider, Walter, and Richard Shiffrin (1977) Controlled and automated human information processing, I: Detection, search and attention. *Psychological Review* 44: 627–44.

Schreyer, Rüdiger (1985) The origin of language: a scientific approach to the study of man. *Topoi* 4: 181–6.

Schuchardt, Hugo (1885) *Über die Lautgesetze*. Berlin: Oppenheim. (Translated by Theo Venneman and Terence H. Wilbur as *On Sound Laws: Against the Neogrammarians*, pp. 1–114. Frankfurt: Athenäum, 1972.)

Schumacher, Helmut, Jacqueline Kubczak, Renate Schmidt, and Vera de Ruiter (2004) *VALBU—Valenzwörterbuch deutscher Verben*. Tübingen: Narr.

Schumann, John H. (1976) Second language acquisition: the pidginization hypothesis. *Language Learning* 26: 391–408.

Schumann, John H. (1978) *The Pidginization Process: A Model for Second Language Acquisition*. Rowley, MA: Newbury House.

Schumann, John H. (1985) Non-syntactic speech in Spanish–English basilang. In R. Andersen (ed.), *Second Language Acquisition: A Cross-Language Perspective*, pp. 355–74. Rowley, MA: Newbury House.

Schütze, Hinrich, Michael Walsh, Travis Wade, and Bernd Möbius (2007) Towards a unified exemplar-theoretic model of phonetic and syntactic phenomena. *Proceedings CogSci 2007* (Nashville).

Schweikert, Walter (2005) *The Order of Prepositional Phrases in the Structure of the Clause*. Amsterdam: Benjamins.

Scollon, Ronald (1976) *Conversations with a One Year Old: A Case Study of the Developmental Foundation of Syntax*. Honolulu: University of Hawaii Press.

Scott, Gary-John (2002) Stacked adjectival modification and the structure of nominal phrases. In G. Cinque (ed.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures*, vol. 1, pp. 91–120. New York: Oxford University Press.

Searle, John R. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Seginer, Yoav (2007) Fast Unsupervised incremental parsing. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics* (ACL 2007), pp. 384–91.

Sekerina, Irina, Karin Stromswold, and Arild Hestvik (2004) How do adults and children process referentially ambiguous pronouns? *Journal of Child Language* 3: 123–52.

Selinker, Larry (1972) Interlanguage. *International Review of Applied Linguistics* 10: 209–31.

Selkirk, Elisabeth (1978) On prosodic structure and its relation to syntactic structure. In T. Fretheim (ed.), *Nordic Prosody*, ed. T. Fretheim, vol. 2, pp. 111–40. Trondheim: TAPIR.

Selkirk, Elisabeth (1982) *The Syntax of Words*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth (1995) The prosodic structure of function words. In Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds.), *University of Massachusetts Occasional Papers: Papers in Optimality Theory*, pp. 439–70. Amherst, MA: GLSA.

Selkirk, Elisabeth, and Tong Shen (1990) Prosodic domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec (eds.), *The Phonology–Syntax Connection*, pp. 313–37. Chicago: University of Chicago Press.

Sells, Peter (ed.) (2001a) *Formal and Empirical Issues in Optimality-Theoretic Syntax*. Stanford, CA: CSLI Publications.

Sells, Peter (2001b) *Structure, Alignment and Optimality in Swedish*. Stanford, CA: CSLI Publications.

Sgall, Petr, Eva Hajicová, and Jarmila Panevova (1986) *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Prague: Academia.

Shadmehr, Reza, and Thomas Brashers-Krug (1997) Functional stages in the formation of human long-term motor memory. *Journal of Neuroscience* 17: 409–19.

Shaffer, Barbara (2000) A syntactic, pragmatic analysis of the expression of necessity and possibility in American Sign Language. Dissertation, University of New Mexico, Albuquerque, NM.

Shaffer, Barbara (2002) CAN T: the negation of modal notions in ASL. *Sign Language Studies* 3: 34–53.

Shieber, Stuart (1986) *An Introduction to Unification-Based Approaches to Grammar*. Stanford, CA: CSLI Publications.

Shieber, Stuart, and Yves Schabes (1991) Generation and synchronous tree adjoining grammars. *Journal of Computational Intelligence* 7: 200–228.

Shimojo, M. (1995) Focus structure and morphosyntax in Japanese: *wa* and *ga*, and word order flexibility. PhD dissertation, SUNY at Buffalo. Available on RRG website: http://linguistics.buffalo.edu/research/rrg.html

Shimojo, M. (2005) *Argument Encoding in Japanese Conversation*. Basingstoke: Palgrave Macmillan.

Shimojo, M. (2008) How missing is the missing verb? The verb-less numeral quantifier construction in Japanese. In Van Valin (2008b: 285–304).

Shimojo, M. (2009) Focus structure and beyond: discourse-pragmatics in RRG. In Guerrero et al. (2009).

Shlonsky, Ur (1997) *Clause Structure and Word Order in Hebrew and Arabic*. New York: Oxford University Press.

Shlonsky, Ur (2000) Subject positions and copular constructions. In H. Bennis, M. Everaert, and E. Reuland (eds.), *Interface Strategies*, pp. 325–47. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Shopen, Timothy (ed.) (2007) *Language Typology and Syntactic Description*, 2nd edn (3 vols). Cambridge: Cambridge University Press.

Sibly, Anne (forthcoming) The semantics of verbs of human physical contact. Honours thesis, Australian National University.

Sigurðsson, Halldór Á. (2000) The locus of case and agreement. *Working Papers in Scandinavian Syntax* 65: 65–108.

Silverstein, M. (1976) Hierarchy of features and ergativity. In R. Dixon (ed.), *Grammatical Categories in Australian Languages*, pp. 112–71. Canberra: Australian Institute of Aboriginal Studies.

Sima'an, Khalil, Alon Itai, Yoad Winter, Alon Altman, and Noa Nativ (2001) Building a tree-bank of modern Hebrew text. *Journal de Traitement Automatique des Langues*, special issue on natural language processing and corpus linguistics 42: 347–80.

Simon, Herbert A. (1974) How big is a chunk? *Science* 183.4124: 482–8.

Simpson, Jane (1983) Aspects of Warlpiri morphology and syntax. PhD thesis, MIT.

Simpson, Jane (1991) *Warlpiri Morpho-Syntax: A Lexicalist Approach*. Dordrecht: Kluwer.

Sinclair, John (1991a) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, John (1991b) Shared knowledge. In *Georgetown University Round Table on Languages and Linguistics 1991*, pp. 489–500. Washington, DC: Georgetown University Press.

Sinclair, John (2001) Review of Biber et al. (1999). *International Journal of Corpus Linguistics* 6: 339–59.

Skinner, Burrhus F. (1957) *Verbal Behavior*. New York: Appleton-Century-Crofts.

Skipper, Jeremy, and Steven Small (2006) fMRI Studies of Language. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, pp. 496–511. Oxford: Elsevier.

Skorik, Pjotr J. (1961/1977) *Grammatika čukotskogo jazyka* [Grammar of Chukchee] (2 vols). Moscow: Izdatel'stvo Akademii Nauk.

Skousen, Royal (1989) *Analogical Modeling of Language*. Dordrecht: Kluwer.

Slobin, Dan I. (1991) Learning to think for speaking: native language, cognition, and rhetorical style. *Pragmatics* 1: 7–25.

Slobin, Dan I. (1996) From 'thought and language' to 'thinking for speaking'. In John J. Gumperz and Stephen C. Levinson (eds.), *Rethinking Linguistic Relativity*, pp. 70–96. Cambridge: Cambridge University Press.

Slobin, Dan I. (2000) Verbalized events: a dynamic approach to linguistic relativity and determinism. In S. Niemeier and R. Dirven (eds.), *Evidence for Linguistic Relativity*, pp. 107–38. Amsterdam: Benjamins.

Slobin, Dan I. (2003) Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner and S. Goldin-Meadow (eds.), *Language in Mind*, pp. 157–91. Cambridge, MA: MIT Press.

Slobin, Dan I. (2005) Issues of linguistic typology in the study of sign language development of deaf children. In Brenda Schick, Marc Marschark, and Patricia E. Spencer (eds.), *Advances in the Sign Language Development of Deaf Children*, pp. 20–45. Oxford: Oxford University Press.

Slobin, Dan I. (2008) Breaking the molds: signed languages and the nature of human language. *Sign Language Studies* 8: 114–30.

Slobin, Dan I., Melissa Bowerman, Penelope Brown, Sonja Eisenbeiss, and Bhuvana Narasimhan (forthcoming) Putting things in places: developmental consequences of linguistic typology. In J. Bohnemeyer and E. Pederson (eds.), *Event Representation*. Cambridge: Cambridge University Press.

Smit, Niels (2005) Noun incorporation in Functional Discourse Grammar. In Casper de Groot and Kees Hengeveld (eds.), *Morphosyntactic Expression in Functional Grammar*, pp. 87–134. Berlin: Mouton de Gruyter.

Smith, C. (1997) *The Parameter of Aspect*, 2nd edn. Dordrecht: Reidel.

Smith, Jennifer L. (2002) Phonological augmentation in prominent positions. Doctoral dissertation, University of Massachusetts, MA.

Smith, Jennifer L. (2007) Representational complexity in syllable structure and its consequences for Gen and Con. In Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.), *Freedom of Analysis?* Berlin: Mouton de Gruyter.

Smith, Marion V. (1996) Linguistic relativity: on hypotheses and confusions. *Communication and Cognition* 29: 65–90.

Smith, Neil (1990) Observations on the pragmatics of tense. *UCL Working Papers in Linguistics* 2: 82–94.

Smith, Neil, and Amahl Smith (1988) A relevance-theoretic account of conditionals. In L. Hyman and C. Li (eds.), *Language, Speech and Mind: Essays in Honor of Victoria Fromkin*, pp. 322–52. London: Routledge.

Smolensky, Paul, Géraldine Legendre, and Bruce Tesar (2006) Optimality Theory: the structure, use, and acquisition of grammatical knowledge. In Paul Smolensky and Geraldine Legendre (eds.), *The Harmonic Mind*, pp. 453–535. Cambridge, MA: MIT Press.

Snider, Neal (2008) An exemplar model of syntactic priming. PhD thesis, Stanford University.

Snyder, William (2007) *Child Language*. Oxford: Oxford University Press.

Sobin, Nicolas (1997) Agreement, default rules, and grammatical viruses. *Linguistic Inquiry* 28: 318–43.

Solan, D., D. Horn, E. Ruppin, and S. Edelman (2005) Unsupervised learning of natural languages. *Proceedings of the National Academy of Science* 102: 11629–34.

Somers, Harold L. (1987) *Valency and Case in Computational Linguistics*. Edinburgh: Edinburgh University Press.

Spencer, Andrew, and Arnold Zwicky (1988a) Introduction. In Spencer and Zwicky (1988b: 1–10).

Spencer, Andrew, and Arnold Zwicky (eds.) (1988b) *The Handbook of Morphology*. Oxford: Blackwell.

Sperber, Dan, and Deirdre Wilson (1986) *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, Dan, and Deirdre Wilson (1990) Spontaneous deduction and mutual knowledge. *Behavioral and Brain Sciences* 13: 179–84.

Sperber, Dan, and Deirdre Wilson (1995) *Relevance: Communication and Cognition*, 2nd edn. Oxford: Blackwell.

Sperber, Dan, and Deirdre Wilson (2002) Pragmatics, modularity and mind-reading. *Mind and Language* 17: 3–23.

Sportiche, Dominique (1988) A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425–50.

Sportiche, Dominique (2007) Division of labor between Merge and Move: strict locality of selection and apparent reconstruction paradoxes. MS, University of California, Los Angeles.

Sproat, Richard (1985) Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3: 173–216.

Squire, Larry R. (1987) *Memory and Brain*. Oxford: Oxford University Press.

Srioutai, Jiranthara (2004) The Thai c1a: a marker of tense or modality? In E. Daskalaki et al. (eds.), *Second CamLing Proceedings* (University of Cambridge), pp. 273–80.

Srioutai, Jiranthara (2006) Time conceptualization in Thai with special reference to d1ay1II, kh3oe:y, k1aml3ang, y3u:I and c1a. PhD thesis. University of Cambridge.

Stampe, David (1973) A dissertation on natural phonology. Doctoral dissertation, University of Chicago. Published by Garland, New York, 1979.

Starke, Michal (2001) Move dissolves into Merge. PhD dissertation, University of Geneva.

Starosta, Stanley (1988) *The Case for Lexicase: An Outline of Lexicase Grammatical Theory*. London: Pinter.

Steedman, Mark (2000) *The Syntactic Process*. Cambridge, MA/London: MIT Press/ Bradford.

Steels, Luc, and Joachim De Beule (2006) A (very) brief introduction to Fluid Construction Grammar. Paper presented at the Third International Workshop on Scalable Natural Language Understanding (ScaNaLU 2006, New York).

Steiner, Erich, and Colin Yallop (2001)*Exploring Translation and Multilingual Text Production: Beyond Content*. Berlin: Mouton de Gruyter.

Steriade, Donca (1988) Reduplication and syllable transfer in Sanskrit and elsewhere. *Phonology* 5: 73–155.

Steriade, Donca (2001) The phonology of perceptibility effects: the P-map and its consequences for constraint organization. MS, MIT.

Stewart, Osamuyimen (2001) *The Serial Verb Construction Parameter*. New York: Garland.

Stokoe, W. C. (1960) *Sign Language Structure*. Silver Spring, MD: Linstok Press.

Stokoe, W. C. (1974) Motor signs as the first form of language. In R. W. Wescott, G. W. Hewes, and W. C. Stokoe (eds.), *Language Origins*, pp. 35–50. Silver Spring, MD: Linkstok Press.

Stokoe, William C. (1980) Sign language structure. *Annual Review of Anthropology* 9: 365–470.

Stokoe, William C. (1991) Semantic phonology. *Sign Language Studies* 71: 107–14.

Stokoe, William C., Dorothy Casterline, and Carl Croneberg (1965) *A Dictionary of American Sign Language on Linguistic Principles*. Washington, DC: Gallaudet College Press.

Strawson, P. F. (1950) On referring. *Mind* 59: 235, 320–44. Reprinted in A. P. Martinich (ed.) *The Philosophy of Language*, 4th edn. Oxford: Oxford University Press, 2001.

Stubbs, M. (1993) British traditions in text analysis: from Firth to Sinclair. In M. Baker, G. Francis, and E. Tognini-Bonelli (eds.), *Text and Technology: In honour of John Sinclair*, pp. 1–33. Amsterdam: Benjamins.

Stump, Gregory (1993) On rules of referral. *Language* 69: 449–79.

Sturt, Patrick (2003) The time-course of the application of Binding Constraints in reference resolution. *Journal of Memory and Language* 48: 542–62.

Stutzman, Verna (1997) A study of the Lou verb phrase. MA dissertation, University of Manitoba. http://mspace.lib.umanitoba.ca/handle/1993/1324

Subirats, Carlos (2007) Relaciones semánticas entre marcos en FrameNet Español. In Juan Cuartero Martina Emsel (ed.), *Vernetzungen. Bedeutung in Wort, Satz und Text. Festschrift für Gerd Wotjak zum 65. Geburtstag*, pp. 357–66. Frankfurt: Lang.

Subirats, Carlos, and Hiroaki Sato (2004) Spanish FrameNet and FrameSQL. *Proceedings of LREC 2004 (BLRSAC)*, pp. 13–16.

Supalla, Ted, and Elissa L. Newport (1978) How many seats in a chair? In Patricia Siple (ed.), *Understanding Language through Sign Language Research*, pp. 91–132. New York: Academic Press.

Suppes, Patrick (1970) Probabilistic grammars for natural languages. *Synthese* 22: 95–116.

Svenonius, Peter (2008a) The position of adjectives and other phrasal modifiers in the decomposition of DP. In L. McNally and C. Kennedy (eds.), *Adjectives and Adverbs: Syntax, Semantics, and Discourse*. Oxford: Oxford University Press, 16–42.

Svenonius, Peter (2008b) Projections of P. In A. Asbury, J. Dotlačil, B. Gehrke, and R. Nouwen (eds.), *Syntax and Semantics of Spatial P*, pp. 63–84. Amsterdam: Benjamins. (http://ling.auf.net/lingBuzz/)

Swart, Henriëtte de, Yoad Winter, and Joost Zwarts (2005) Bare predicate nominals in Dutch. In Emar Maier, Corien Bary, and Janneke Huitink (eds.), *Proceedings of SuB 9*, pp. 446–60.

Swart, Henriëtte de, Yoad Winter, and Joost Zwarts (2007) Bare nominals and reference to capacities. *Natural Language and Linguistic Theory* 25: 195–222.

Swart, Henriëtte de, and Joost Zwarts (2009) Less form–more meaning: why bare singular nouns are special. *Lingua* 119: 280–95.

Sweetser, Eve E. (1990) *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Swinney, David A. (1979) Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18: 645–59.

Swinney, David A., Marilyn Ford, Ili Frauenfelder, and Joan Bresnan (1988) On the temporal course of gap-filling and antecedent assignment during sentence comprehension. In B. Grosz, R. Kaplan, and M. Macken (eds.), *Language Structure and Processing*. Stanford, CA: CILT Publications.

Sysoeva, Anna, and Katarzyna M. Jaszczolt (2007) Composing utterance meaning: an interface between pragmatics and psychology. Paper presented at the 10th International Pragmatics Conference (Göteborg).

Sysoeva, Anna, and Katarzyna M. Jaszczolt (2009) More than radical pragmatics: primary meaning without the syntactic constraint. Unpublished paper, University of Cambridge.

Talmy, Leonard (1972) Semantic structures in English and Atsugewi. PhD dissertation, University of California, Berkeley.

Talmy, Leonard (1983) How languages structure space. In H. Pick and L. Acredolo (eds.), *Spatial Orientation: Theory, Research and Application*, pp. 225–82. New York: Plenum Press.

Talmy, Leonard (1985) Lexicalization patterns: semantic structure in lexical form. In T. Shopen (ed.), *Grammatical Categories and the Lexicon*, vol. 3, pp. 57–149. Cambridge: Cambridge University Press.

Talmy, Leonard (1988) The relation of grammar to cognition. In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, pp. 165–205. Amsterdam: Benjamins.

Talmy, Leonard (1996) Fictive motion in language and 'ception'. In Paul Bloom, M. A. Peterson, L. Nadel, and M. F. Garrett (eds.), *Language and Space*, pp. 211–76. Cambridge, MA/London: MIT Press/Bradford.

Talmy, Leonard (2000) *Toward a Cognitive Semantics*, 1: *Concept Structuring Systems*. Cambridge, MA: MIT Press.

Talmy, Leonard (2008) The relation of grammar to cognition. In Vyvyan Evans, Benjamin Bergen, and Jörg Zinken (eds.), *The Cognitive Linguistics Reader*, pp. 381–544. London: Equinox.

Tanaka, Keiko (1994) *Advertising Language: A Pragmatic Approach to Advertisements in Britain and Japan*: London: Routledge.

Tannen, Deborah (1993) *Framing in Discourse*. Oxford University Press.

Taylor, John R. (1989) *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Clarendon.

Taylor, John R. (2004) *Linguistic Categorization: Prototypes in Linguistic Theory*, 3rd edn. Oxford: Clarendon Press.

Teich, Elke (1999) *Systemic Functional Grammar in Natural Language Generation: Linguistic Description and Computational Representation*. London: Cassell.

Tenbrink, Thora (2007) *Space, Time, and the Use of Language: An Investigation of Relationships*. Berlin: Mouton de Gruyter.

Tenenbaum, Joshua, Thomas Griffiths, and Charles Kemp (2006) Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences* 10: 309–18.

Tenpenny, Patricia L. (1995) Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin and Review* 2: 339–63.

Terrace, Herbert S. (1985) In the beginning was the 'name'. *American Psychologist* 40: 1011–28.

Teruya, K., E. Akerejola, T. Andersen, A. Caffarel, J. Lavid, C. Matthiessen, U. H. Petersen, P. Patpong, and F. Smedegaard (2007) Typology of MOOD: a text-based and system-based functional view. In Christian M. I. M. Matthiessen and Jonathan Webster (eds.), *Continuing Discourse on Language: A Functional Perspective*, vol. 2. Equinox: London.

Tesar, Bruce (1998) Using the mutual inconsistency of structural descriptions to overcome ambiguity in language learning. In Pius N. Tamanji and Kiyomi Kusumoto (eds.), *Proceedings of NELS 28*, pp. 469–83. Amherst, MA: GLSA Publications.

TESAR, BRUCE, JANE GRIMSHAW, and ALAN PRINCE (1999) Linguistic and cognitive explanation in Optimality Theory. In Ernest Lepore and Zenon Pylyshyn (eds.), *What is Cognitive Science?*, pp. 295–326. Oxford: Blackwell.

TESAR, BRUCE, and PAUL SMOLENSKY (2000) *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

TESNIÈRE, LUCIEN (1953) *Esquisse d'une syntaxe structurale*. Paris: Klincksieck.

TESNIÈRE, LUCIEN (1976) *Éléments de syntaxe structurale*, 2nd rev. and corrected edn. Paris: Klincksieck.

TESNIÈRE, LUCIEN (1980) *Grundzüge der strukturalen Syntax*, ed. and trans. Ulrich Engel. Stuttgart: Klett-Cotta.

TETTAMANTI, MARCO, GIOVANNI BUCCINO, MARIA C. SACCUMAN, VITTORIO GALLESE, MASSIMO DANNA, PAOLA SCIFO, FERRUCCIO FAZIO, GIACOMO RIZZOLATTI, STEFANO F. CAPPA, and DANIELA PERANI (2005) Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience* 17: 273–81.

TEUBERT, WOLFGANG (2003) Die Valenz nichtverbaler Wortarten: das Substantiv. In Ágel et al. (2003: 820–35).

THOMASON, RICHMOND H. (1997) Nonmonotonicity in linguistics. In J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*, pp. 777–831. Oxford: Elsevier Science.

THOMPSON, SANDRA A. (1983) Grammar and discourse: the English detached participial clause. In F. KleinAndreu (ed.), *Discourse Perspectives on Syntax*. New York: Academic Press, pp. 43–65.

THOMPSON, SANDRA A. (1985) Grammar and written discourse: initial vs. final purpose clauses in English. *Text* 5: 55–84.

THOMPSON, SANDRA A., and PAUL J. HOPPER (2001) Transitivity, clause structure, and argument structure: evidence from conversation. In Bybee and Hopper (2001: 27–60).

THOMPSON, SANDRA A., and ANTHONY MULAC (1991a) The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics* 15: 237–51.

THOMPSON, SANDRA A., and ANTHONY MULAC (1991b) A quantitative perspective on the grammaticization of epistemic parentheticals in English. In Traugott and Heine (1991b: 313–29).

THRÁINSSON, HÖSKULDUR (1996) On the (non)-universality of functional categories. In W. Abraham, S. D. Epstein, H. Thráinsson, and C. J.-W. Zwart (eds.), *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, pp. 253–81. Amsterdam: Benjamins.

TIERSMA, PETER (1982) Local and general markedness. *Language* 58: 832–49.

TOGNINI-BONELLI, ELENA (2001) *Corpus Linguistics at Work*. Amsterdam: Benjamins.

TOIVONEN, IDA (2001) The phrase structure of non-projecting words. PhD thesis, Stanford University.

TOIVONEN, IDA (2003) *Non-projecting Words: A Case Study of Swedish Verbal Particles*. Dordrecht: Kluwer.

TOMASELLO, M. (2000) Do young children have adult syntactic competence? *Cognition* 74: 209–53.

TOMASELLO, MICHAEL (2003) *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

TOMASELLO, MICHAEL, and J. CALL (1997) *Primate Cognition*. Oxford: Oxford University Press.

Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll (2005) Understanding and sharing intentions: the origins of cultural cognition. *Brain and Behavioral Science* 28: 675–91.

Tomlin, Russell S. (1995) Focal attention, voice, and word order. In Pamela Downing and Michael Noonan (eds.), *Word Order in Discourse*, pp. 517–54. Amsterdam: Benjamins.

Torres-Cacoullos, Rena (2001) From lexical to grammatical to social meaning. *Language in Society* 30: 443–78.

Tottie, Gunnel (1991a) *Negation in English Speech and Writing: A Study in Variation*. San Diego, CA: Academic Press.

Tottie, Gunnel (1991b) Lexical diffusion in syntactic change: frequency as a determinant of linguistic conservatism in the development of negation in English. In Dieter Kastovsky (ed.), *Historical English Syntax*, pp. 439–67. Berlin: Mouton de Gruyter.

Touretzky, David (1986) *The Mathematics of Inheritance Systems*. Los Altos, CA: Morgan Kaufmann.

Townsend, Charles, and Laura Janda (1996) *Common and Comparative Slavic: Phonology and Inflection*. Columbus, OH: Slavica.

Tracy-Ventura, Nicole, Douglas Biber, and Viviana Cortes (2007) Lexical bundles in Spanish speech and writing. In G. Parodi (ed.), *Working with Spanish Corpora*, pp. 217–31. London: Continuum.

Traugott, Elizabeth C. (1980) Meaning-change in the development of grammatical markers. *Language Sciences* 2: 44–61.

Traugott, Elizabeth C. (1989) On the rise of epistemic meaning: an example of subjectification in semantic change. *Language* 65: 31–55.

Traugott, Elizabeth C. (2004a) Historical pragmatics. In Horn and Ward (2004: 538–61).

Traugott, Elizabeth C. (2004b) A critique of Levinson's view of Q- and M-inferences in historical pragmatics. *Journal of Historical Pragmatics* 5: 1–25.

Traugott, Elizabeth C. (2006) Historical aspects of modality. In W. Frawley (ed.), *The Expression of Modality*, pp. 107–39. Berlin: Mouton de Gruyter.

Traugott, Elizabeth C., and Richard B. Dasher (2002) *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Traugott, Elizabeth C., and Bernd Heine (eds.) (1991a) *Approaches to Grammaticalization*, vol. 1. Amsterdam: Benjamins.

Traugott, Elizabeth C., and Bernd Heine (eds.) (1991b) *Approaches to Grammaticalization*, vol. 2. Amsterdam: Benjamins.

Traugott, Elizabeth C., and Ekkehard König (1991) The semantics-pragmatics of grammaticalization revisited. In Traugott and Heine (1991a: 189–218). Amsterdam: Benjamins.

Travis, Catherine (2004) The ethnopragmatics of the diminutive in conversational Colombian Spanish. *Intercultural Pragmatics* 1: 249–74.

Travis, Catherine (2006) The communicative realisation of *confianza* and *calor humano* in Colombian Spanish. In Cliff Goddard (ed.), *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 199–229. Berlin: Mouton de Gruyter. Pp. 199–229.

Travis, Lisa (1984) Parameters and effects of word order variation. PhD thesis, MIT.

Traxler, Matthew, and Martin Pickering (1996) Plausibility and the processing of unbounded dependencies: an eye-tracking study. *Journal of Memory and Language* 35: 454–75.

Tsai, Dylan (2007) *Left Periphery and Why–How Alternations*. Taiwan: National Tsing-Hua University.

Tsujimura, Natsuko (2006) *An Introduction to Japanese Linguistics*. Oxford: Blackwell.

Tucker, D. M. (1992) Developing emotions and cortical networks. In M. Gunnar and C. Nelson (eds.), *Developmental Neuroscience: Minnesota Symposium on Child Psychology 24*, pp. 75–128. Hillsdale, NJ: Erlbaum.

Tuggy, David (2007) Schematicity. In Dirk Geeraerts and Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, pp. 82–116. Oxford: Oxford University Press.

Ullman, Michael (2001) The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition* 4: 105–22.

Ungerleider, Leslie G., and Mortimer Mishkin (1982) Two cortical visual systems. In D. G. Ingle, M. A. Goodale, and R. J. Q. Mansfield (eds.), *Analysis of Visual Behavior*, pp. 549–86. Cambridge: MIT Press.

Uriagereka, Juan (1998) *Rhyme and Reason*. Cambridge, MA: MIT Press.

Uriagereka, Juan (1999) Multiple spell-out. In Samuel D. Epstein and Norbert Hornstein (eds.), *Working Minimalism*, pp. 251–82. Cambridge, MA: MIT Press.

Uriagereka, Juan (2008) *Syntactic Anchors*. Cambridge: Cambridge University Press.

Uzonyí, Pál (2003) Dependenzstruktur und Konstituenzstruktur. In Ágel et al. (2003: 230–47).

Van Craenenbroeck, Jeroen (2006) Transitivity failures in the left periphery and foot-driven movement operations. *Linguistics in the Netherlands* 23: 52–64.

van der Auwera, Johan, and Vladimir A. Plungian (1998) Modality's semantic map. *Linguistic Typology* 2: 79–124.

Van der Henst, Jean, and Dan Sperber (2004) Testing the cognitive and the communicative principles of relevance. In I. Noveck and D. Sperber (eds.), *Experimental Pragmatics*, pp. 229–79. London: Palgrave.

Vanderveken, Daniel (2002) Universal grammar and speech act theory. In D. Vanderveken and S. Kubo (eds.), *Essays in Speech Act Theory*, pp. 25–62. Amsterdam: Benjamins.

van Hoek, Karen (1995) Conceptual reference points: a cognitive grammar account of pronominal anaphora constraints. *Language* 71: 310–40.

van Hoek, Karen (1997) *Anaphora and Conceptual Structure*. Chicago: University of Chicago Press.

Van Overschelde, James P., Katherine A. Rawson, and John Dunlosky (2004) Category norms: an updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language* 50: 289–335.

Van Valin, R. D., Jr. (1977a) Ergativity and the universality of subjects. *CLS* 13: 689–706.

Van Valin, R. D., Jr. (1977b) Aspects of Lakhota syntax. PhD dissertation, University of California, Berkeley.

Van Valin, R. D., Jr. (1981) Grammatical relations in ergative languages. *Studies in Language* 5: 361–94.

Van Valin, R. D., Jr. (1985) Case marking and the structure of the Lakhota clause. In J. Nichols and A. Woodbury (eds.), *Grammar Inside and Outside the Clause*, pp. 363–413. Cambridge: Cambridge University Press.

Van Valin, R. D., Jr. (ed.) (1993a) *Advances in Role and Reference Grammar*. Amsterdam: Benjamins.

Van Valin, R. D., Jr. (1993b) A synopsis of Role and Reference Grammar. In Van Valin (1993a), pp. 1–164.

Van Valin, R. D. Jr (1994) Extraction restrictions, competing theories and the argument from the poverty of the stimulus. In S. Lima, R. Corrigan, and G. Iverson (eds.), *The Reality of Linguistic Rules*, pp. 243–59. Amsterdam: Benjamins.

Van Valin, R. D., Jr (1995) Toward a functionalist account of so-called 'extraction' constraints. In B. Devriendt, L. Goossens, and J. van der Auwera (eds.), *Complex Structures: A Functionalist Perspective*, pp. 29–60. Berlin: Mouton de Gruyter.

Van Valin, R. D., Jr. (1998) The acquisition of WH-questions and the mechanisms of language acquisition. In M. Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, pp. 221–49. Mahwah, NJ: Erlbaum.

Van Valin, R. D., Jr. (1999) Generalized semantic roles and the syntax–semantics interface. In F. Corblin, C. Dobrovie-Sorin, and J.-M. Marandin (eds.), *Empirical Issues in Formal Syntax and Semantics 2*, pp. 373–89. The Hague: Thesus.

Van Valin, R. D., Jr. (2000) Functional linguistics. In Aronoff and Rees-Miller (2000: 319–36).

Van Valin, R. D., Jr. (2001a) *An Introduction to Syntax*. Cambridge: Cambridge University Press.

Van Valin, R. D., Jr (2001b) The acquisition of complex sentences: a case study in the role of theory in the study of language development. *CLS* 36.2: 511–31.

Van Valin, R. D., Jr. (2002) The development of subject–auxiliary inversion in English wh-questions: an alternative analysis. *Journal of Child Language* 29: 161–75.

Van Valin, R. D., Jr. (2004) Semantic macroroles in Role and Reference Grammar. In R. Kailuweit and M. Hummel (eds.), *Semantische Rollen*, pp. 62–82. Tübingen: Narr.

Van Valin, R. D., Jr. (2005) *Exploring the Syntax–Semantics Interface*. Cambridge: Cambridge University Press.

Van Valin, R. D., Jr. (2006) Semantic macroroles and language processing. In Bornkessel (2006: 263–302).

Van Valin, R. D., Jr. (2007) The Role and Reference Grammar analysis of three-place predicates. *Suvremena lingvistika* 33: 31–63.

Van Valin, R. D., Jr. (2008a) RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In Van Valin (2008a: 161–78).

Van Valin, R. D., Jr. (ed.) (2008b) *Investigations of the Syntax–Semantic–Pragmatics Interface.* Amsterdam: Benjamins.

Van Valin, R. D., Jr. (2009a) Privileged syntactic arguments, pivots, and controllers. In Guerrero et al. (2009).

Van Valin, R. D., Jr. (2009b) Role and Reference Grammar. In Bernd Heine and Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, pp. 703–38.

Van Valin, R. D., Jr. (forthcoming a) Lexical representation, co-composition, and linking syntax and semantics. In J. Pustejovsky and P. Bouillon (eds.), *New Developments in the Generative Lexicon*. Dordrecht: Kluwer.

Van Valin, R. D., Jr., and Randy LaPolla (1997) *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Van Valin, R. D., Jr., and David Wilkins (1993) Predicting syntactic structure from semantic representations: *remember* in English and its equivalents in Mparntwe Arrernte. In Van Valin (1993a: 499–534).

Van Valin R. D., Jr., and David Wilkins (1996) The case for effector: case roles, agents and agency revisited. In M. Shibatani and S. Thompson (eds.), *Grammatical Constructions*, pp. 289–322. Oxford: Oxford University Press.

van Zaanen, Menno (2000) ABL: Alignment-based learning. *Proceedings COLING 2000* (Saarbrücken).

Venditti, Jennifer J. (2005) The J_ToBI model of Japanese intonation. In Sun-Ah Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, pp. 172–200. Oxford: Oxford University Press.

Vendler, Z. (1967) *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.

Vihman, Marilyn May (1980) Sound change and child language. In Elizabeth Traugott, Roberta Labrum, and Susan Shepherd (eds.), *Papers from the 4th International Conference on Historical Linguistics*, pp. 303–20. Amsterdam: Benjamins.

Viertel, John (1966) Concepts of language underlying the eighteenth century controversy about the origins of language. *Georgetown Monograph Series on Language and Linguistics* 19: 109–32.

Vincent, Nigel (2001) LFG as a model of syntactic change. In Butt and King (2001b: 1–42).

Voeltz, F. K. Erhard, and Christa Kilian-Hatz (eds.) (2001) *Ideophones*. Amsterdam: Benjamins.

von Heusinger, K. and K. Turner (eds.) (2006) *Where Semantics Meets Pragmatics*. Oxford: Elsevier.

Vries, Lourens de (1985) Topic and focus in Wambon discourse. In A. Machtelt Bolkestein, Casper de Groot, and J. Lachlan Mackenzie (eds.), *Syntax and Pragmatics in Functional Grammar*, pp. 155–80. Dordrecht: Foris.

Vygotsky, L. S. (1986) *Thought and Language (Myshlenie i Rech)*. Cambridge, MA: MIT Press.

Walker, Richard (1979) Central carrier phonemics. In D. Zimmerly (ed.), *Contributions to Canadian Linguistics*, pp. 93–107. Ottawa: National Museums of Canada.

WALS: see Haspelmath et al. (2005).

Walton, C. (1986) *Sama Verbal Semantics: Classification, Derivation and Inflection*. Manila: Linguistic Society of the Philippines.

Wanders, Gerry (forthcoming) Typology and iachrony: on the use of the subjunctive in adverbial clauses in the Ibero-Romance languages.

Ward, G. L. (1990) The discourse functions of VP preposing. *Language* 66: 742–63.

Wasow, Thomas (2002) *Postverbal Behavior.* Stanford, CA: CSLI Publications.

Watson, Helen (1987) Learning to apply numbers to nature. *Educational Studies in Mathematics* 18: 339–57.

Watters, John R. (1979) Focus in Aghem: a study of its formal correlates and typology. In Larry Hyman (ed.), *Aghem Grammatical Structure*, pp. 137–97. Los Angeles: University of Southern California.

Watters, J. (1993) An investigation of Turkish clause linkage. In Van Valin (1993a: 535–60).

Way, Andy (1999) A hybrid architecture for robust MT using LFG-DOP. Special issue on memory-based learning, *Journal of Experimental and Theoretical Artificial Intelligence* 11: 441–71.

Weber, Heinz J. (1997) *Dependenzgrammatik: Ein interaktives Arbeitsbuch*, 2nd edn. Tübingen: Narr.

Wedel, Andrew (2006) Exemplar models, evolution and language change. *Linguistic Review* 23: 24–74.

Wei, Li (2006) Bilingualism. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, pp. 1–12. Oxford: Elsevier.

Weinberg, Steven (1976) The forces of nature. *Bulletin of the American Academy of Arts and Sciences* 29: 13–29.

Weist, R. M. (2002) The first language acquisition of tense and aspect: a review. In R. Salaberry and Y. Shirai (eds.), *Tense–Aspect Morphology in L2 Acquisition*, pp. 21–78. Amsterdam: Benjamins.

Weist, R., A. Pawlak and J. Carapella (2004) Syntactic-semantic interface in the acquisition of verb morphology. *Journal of Child Language* 31: 31–60.

Welke, Klaus M. (1988) *Einführung in die Valenz- und Kasustheorie*. Leipzig: VEB Bibliographisches Institut.

Welke, Klaus M. (1995) Dependenz, Valenz und Konstituenz. In Ludwig M. Eichinger and Hans-Werner Eroms (eds.), *Dependenz und Valenz*, pp. 163–75. Hamburg: Buske.

Welke, Klaus M. (2005) *Deutsche Syntax funktional: Perspektiviertheit syntaktischer Strukturen*, 2nd edn. Tübingen: Stauffenburg.

Welker, Katherine (1994) Plans in the common ground. PhD dissertation. Ohio State University.

Wellman, Henry M. (1990) *The Child's Theory of Mind*. Cambridge, MA: MIT Press.

Wertheimer, Max (1923) Laws of organization in perceptual forms. Translated in W. D. Ellis (ed.), *A Source Book of Gestalt Psychology*, pp. 71–88. London: Routledge & Kegan Paul, 1938.

Wescoat, Michael T. (2002) On lexical sharing. PhD thesis, Stanford University.

Wescoat, Michael T. (2005) English nonsyllabic auxiliary contractions: an analysis in LFG with lexical sharing. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG05 Conference*, pp. 468–86. Stanford, CA: CSLI Publications.

Wescoat, Michael T. (2007) Preposition-determiner contractions: an analysis in Optimality-Theoretic Lexical-Functional Grammar with lexical sharing. In Butt and King (2007: 439–59).

Wexler, Kenneth 2004. Lenneberg's dream. In Lyle Jenkins (ed.), *Variation and Universals in Biolinguistics*, pp. 239–84. Amsterdam: Elsevier.

Wexler, Kenneth, and Yu-Chin Chien (1985) The development of lexical anaphors and pronouns. *Papers and Reports on Child Language Development* 24: 138–49.

White, M. (1994) Language in job interviews: differences relating to success and socioeconomic variables. PhD dissertation, Northern Arizona University.

Whitehead, Alfred N. (1925) *Science and the Modern World*. New York: Free Press.

Whiten, Andrew (ed.) (1991) *Natural Theories of Mind*. Oxford: Blackwell.

Whorf, Benjamin Lee (1956) *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*, edited by John B. Carroll. Cambridge, MA: MIT Press.

Wiebe, Grace, Gary Libben, Tom Priestly, Ron Smyth, and Sam Wang (eds.) (2006) *Phonology, Morphology, and the Empirical Imperative: Papers in Honour of Bruce Derwing*. Taipei: Crane Publishing Company.

Wiemer, Björn (2004) Applications of RRG in diachronic syntax. In Nolan (2004a: 349–53).

Wierzbicka, Anna (1972) *Semantic Primitives*. Frankfurt: Athenaeum.

Wierzbicka, Anna (1979) Ethno-syntax and the philosophy of grammar. *Studies in Language* 3: 313–83.

Wierzbicka, Anna (1980) *The Case for Surface Case*. Ann Arbor: Karoma.

Wierzbicka, Anna (1985) *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.

Wierzbicka, Anna (1988) *The Semantics of Grammar*. Amsterdam: Benjamins.

Wierzbicka, Anna (1992) *Semantics, Culture, and Cognition*. Oxford: Oxford University Press.

Wierzbicka, Anna (1996a) *Semantics: Primes and Universals*. Oxford: Oxford University Press.

Wierzbicka, Anna (1996b) Japanese cultural scripts: cultural psychology and cultural grammar. *Ethos* 24: 527–55.

Wierzbicka, Anna (1997) *Understanding Cultures Through Their Key Words*. Oxford: Oxford University Press.

Wierzbicka, Anna (1998) Anchoring linguistic typology in universal concepts. *Linguistic Typology* 2: 141–94.

Wierzbicka, Anna (1999) *Emotions Across Languages and Cultures*. Cambridge: Cambridge University Press.

Wierzbicka, Anna (2002a) Semantic primes and linguistic typology. In Cliff Goddard and Anna Wierzbicka (eds.), *Meaning and Universal Grammar: Theory and Empirical Findings*, vol. 2, pp. 257–300. Amsterdam: Benjamins.

Wierzbicka, Anna (2002b) English causative constructions in an ethnosyntactic perspective: focusing on LET. In N. J. Enfield (ed.), *Ethnosyntax*, pp. 162–203. Oxford: Oxford University Press.

Wierzbicka, Anna (2003[1991]) *Cross-Cultural Pragmatics: The Semantics of Human Interaction*, 2nd edn. Berlin: Mouton de Gruyter.

Wierzbicka, Anna (2006a) *English: Meaning and Culture*. New York: Oxford University Press.

Wierzbicka, Anna (2006b) Anglo cultural scripts against 'putting pressure' on other people, and their linguistic manifestations. In Cliff Goddard (ed.), *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 31–63. Berlin: Mouton de Gruyter.

Wierzbicka, Anna (2007a) Bodies and their parts: an NSM approach to semantic typology. *Language Sciences* 29: 14–65.

Wierzbicka, Anna (2007b) Shape and colour in language and thought. In Andrea C. Schalley and Drew Khlentzos (eds.), *Mental States*, vol. 2: *Language and Cognitive Structure*, pp. 37–60. Amsterdam: Benjamins.

Wierzbicka, Anna (2008a) Why there are no 'colour universals' in language and thought? *Journal of the Royal Anthropological Institute* 14: 407–25.

Wierzbicka, Anna (2008b) Case in NSM: a reanalysis of the Polish dative. In Andrej Malchukov and Andrew Spencer (eds.), *The Oxford Handbook of Case*, pp. 151–69. Oxford: Oxford University Press.

Wierzbicka, Anna (2009) All people eat and drink: does this mean that 'eat' and 'drink' are universal human concepts? In John Newman (ed.), *The Linguistics of Eating and Drinking*, pp. 65–89. Amsterdam: Benjamins.

Wierzbicka, Anna (forthcoming) *Evidence, Experience, Sense: The Hidden Cultural Legacy of English.* New York: Oxford University Press.

Wilbur, Ronnie B. (1987) *American Sign Language: Linguistic and Applied Dimensions*. Boston, MA: College-Hill Press.

Wilbur, Ronnie B., and Brenda Schick (1987) The effects of linguistic stress on ASL signs. *Language and Speech* 30: 301–23.

Wilcox, Phyllis P. (2000) *Metaphor in American Sign Language*. Washington, DC: Gallaudet University Press.

Wilcox, Sherman (2002) The iconic mapping of space and time in signed languages. In Liliana Albertazzi (ed.), *Unfolding Perceptual Continua*, pp. 255–81. Amsterdam: Benjamins.

Wilcox, Sherman (2004a) Cognitive iconicity: conceptual spaces, meaning, and gesture in signed languages. *Cognitive Linguistics* 15: 119–47.

Wilcox, Sherman (2004b) Gesture and language: cross-linguistic and historical data from signed languages. *Gesture* 4: 43–75.

Wilcox, Sherman (2005) Routes from gesture to language. *Revista da ABRALIN—Associação Brasileira de Lingüística* 4: 11–45.

Wilcox, Sherman (2007) Routes from gesture to language. In Elena Pizzuto Antinori, Paola Pietrandrea, and Raffele Simone (eds.), *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*, pp. 107–31. Berlin: Mouton de Gruyter.

Wilcox, Sherman, and Barbara Shaffer (2006) Modality in American Sign Language. In William Frawley (ed.), *The Expression of Modality*, pp. 207–37. Berlin: Mouton de Gruyter.

Wilcox, Sherman, Barbara Shaffer, Maria J. Jarque, Josep M. S. Valenti, and Elena Pizzuto (2000) The emergence of grammar from word and gesture: a cross-linguistic study of modal verbs in three signed languages. *7th International Conference on Theoretical Issues in Sign Language Research* (Amsterdam).

Wilcox, Sherman, and Phyllis Wilcox (1995) The gestural expression of modality in American Sign Language. In Joan Bybee and Suzanne Fleischman (eds.), *Modality in Grammar and Discourse*, pp. 135–62. Amsterdam: Benjamins.

Wilcox, Sherman, and Phyllis Wilcox (2003) Feeling metonymy and metaphor: evidence from American Sign Language derivational morphology. Paper presented at the International Cognitive Linguistics Conference, Logroño.

Wilcox, Sherman, Phyllis Wilcox, and Maria J. Jarque (2003) Mappings in conceptual space: metonymy, metaphor, and iconicity in two signed languages. *Jezikoslovlje* 4: 139–56.

Wilkins, David (1991) The semantics, pragmatics and diachronic development of 'associated motion' in Mparntwe Arrernte. *Buffalo Papers in Linguistics 91-01*, pp. 207–57. Dept of Linguistics, University at Buffalo.

Willems, Klaas, and Ann Coene (2006) Satzmuster und die Konstruktionalität der Verbbedeutung: Überlegungen zum Verhältnis von Konstruktionsgrammatik und Valenztheorie. *Sprachwissenschaft* 31: 237–72.

Williams, Edwin (1985) PRO and the subject of NP. *Natural Language and Linguistic Theory* 3: 297–315.

Wilson, Colin (2001) Consonant cluster neutralization and targeted constraints. *Phonology* 18: 147–97.

Wilson, Deirdre (1994) Relevance and understanding. In G. Brown, K. Malmkjær, A. Pollit, and J. Williams (eds.), *Language and Understanding*, pp. 35–58. Oxford: Oxford University Press.

Wilson, Deirdre (2002) Relevance: the cognitive principle. In class notes, 'Pragmatic Theory', University College London.

Wilson, Deirdre, and Robyn Carston (2006) Metaphor, relevance and the 'emergent property' issue. *Mind and Language* 21: 404–33.

Wilson, Deirdre, and Robyn Carston (2007) A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In Burton-Roberts (2007: 230–59).

Wilson, Deirdre, and Dan Sperber (1981) On Grice's theory of conversation. In P. Werth (ed.), *Conversation and Discourse*, pp. 155–78. London: Croom Helm.

Wilson, Deirdre, and Dan Sperber (1988) Mood and the analysis of non-declarative sentences. In J. Dancy, J. Moravcsik, and C. Taylor (eds.), *Human Agency: Language, Duty and Value*, pp. 77–101. Stanford, CA: Stanford University Press.

Wilson, Deirdre, and Dan Sperber (1992) On verbal irony. *Lingua* 87: 53–76.

Wilson, Deirdre, and Dan Sperber (2002a) Relevance theory. *UCL Working Papers in Linguistics* 14: 249–90.

Wilson, Deirdre and Sperber, Dan (2002b) Truthfulness and relevance. *Mind* 111: 583–632.

Wilson, Deirdre, and Dan Sperber (2005) Reply to Rajagopalan. *Intercultural Pragmatics* 2: 99–103.

Wilson, Stephen (2003) Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language* 30: 75–115.

Winograd, Terence (1972) *Understanding Natural Language*. New York: Academic Press.

Wittgenstein, Ludwig (1953) *Philosophical Investigations*. Oxford: Blackwell.

Wolfe-Quintero, K., S. Inagaki, and H.-Y. Kim (1998) *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity*. Honolulu: University of Hawaii Press.

Wong, Jock (2004) Reduplication of nominal modifiers in Singapore English. *World Englishes* 23: 339–54.

Wong, Jock (2006) Social hierarchy in the 'speech culture' of Singapore. In Cliff Goddard (ed.), *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 99–125. Berlin: Mouton de Gruyter.

Wong, Jock (2008) Anglo English and Singapore English tags: their meanings and cultural significance. *Pragmatics and Cognition* 16: 88–117.

Wong, Jock, Cliff Goddard, and Anna Wierzbicka (forthcoming) *Walk*, *run*, *jump*: the semantics of human locomotion.

Wong, Yuk, and Raymond Mooney (2007) Learning synchronous grammars for semantic parsing with lambda calculus. *Proceedings ACL-2007*, pp. 960–67.

Woolford, Ellen (2001) Case patterns. In Géraldine Legendre, Jane Grimshaw, and Sten Vikner (eds.) *Optimality-Theoretic Syntax*, pp. 509–43. Cambridge, MA: MIT Press.

Wylie, Laurence, and Rick Stafford (1977) *Beaux Gestes: A Guide to French Body Talk*. Cambridge: Undergraduate Press.

Yang, Charles (2004) Universal Grammar, statistics or both? *Trends in Cognitive Sciences* 8: 451–6.

YE, ZHENGDAO (2004) The Chinese folk model of facial expressions: a linguistic perspective. *Culture and Psychology* 10: 195–222.

YE, ZHENGDAO (2006) Why the inscrutable Chinese face? Emotionality and facial expressions in Chinese. In Cliff Goddard (ed.), *Ethnopragmatics: Understanding Discourse in Cultural Context*, pp. 127–69. Berlin: Mouton de Gruyter.

YE, ZHENGDAO (2007) Memorisation, learning and cultural cognition: the notion of *bèi* ('auditory memorisation') in the written Chinese tradition. In Mengistu Amberber (ed.), *The Language of Memory in a Crosslinguistic Perspective*, pp. 139–80. Amsterdam: Benjamins.

YOON, KYUNG-JOO (2004) Not just words: Korean social models and the use of honorifics. *Intercultural Pragmatics* 1: 189–210.

YOU, LIPING, and KAIYING LIU (2005) Building Chinese FrameNet database. *Natural Language Processing and Knowledge Engineering*: 301–6.

YOU, LIPING, TAO LIU, and KAIYING LIU (2007) Chinese FrameNet and OWL representation. *Sixth International Conference on Advanced Language Processing and Web Information Technology* (ALPIT 2007), pp. 140–45.

YUS, FRANCISCO (1997) Grammar: relevance-theoretic concerns. *Revista Alicantina de Estudios Ingleses* 10: 235–48.

YUS, FRANCISCO (1998a) A decade of relevance theory. *Journal of Pragmatics* 30: 305–45.

YUS, FRANCISCO (1998b) Relevance theory and media discourse: a verbal-visual model of communication. *Poetics* 25: 293–309.

YUS, FRANCISCO (2001) *Ciberpragmática: el uso del lenguaje en Internet*. Barcelona: Ariel.

YUS, FRANCISCO (2003) Humor and the search for relevance. *Journal of Pragmatics* 35: 1295–1331.

YUS, FRANCISCO (2006) Relevance theory. In K. Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn, vol. 10, pp. 512–19.

ZAENEN, ANNIE (1980) Extraction rules in Icelandic. PhD thesis, Harvard University. Reprinted as Zaenen (1985).

ZAENEN, ANNIE (1985) *Extraction Rules in Icelandic*. New York: Garland.

ŽEGARAC, VLADIMIR (1990) Pragmatics and verbal aspect. *UCL Working Papers in Linguistics* 2: 113–43.

ŽEGARAC, VLADIMIR (1993) Some observations on the pragmatics of the progressive. *Lingua* 90: 201–20.

ZESHAN, ULRIKE (2004) Interrogative constructions in signed languages: crosslinguistic perspectives. *Language* 80: 7–39.

ZIEMKE, TOM, JORDAN ZLATEV, and ROSLYN M. FRANK (eds.) (2007) *Body, Language and Mind*, vol. 1: *Embodiment*. Berlin: Mouton de Gruyter.

ZIFONUN, GISELA, LUDGER HOFFMANN, BRUNO STRECKER, et al. (1997) *Grammatik der deutschen Sprache* (3 vols). Berlin: De Gruyter.

ZIPF, GEORGE K. (1949) *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

ZOLLMANN, ANDREAS, and KHALI SIMA'AN. (2005) A consistent and efficient estimator for data-oriented parsing. *Journal of Automata, Languages and Combinatorics* 10: 367–88.

ZUIDEMA, WILLEM (2006) What are the productive units of Natural Language Grammar? A DOP approach to the automatic identification of constructions. *Proceedings CoNLL 2006*, pp. 29–36.

Zuidema, Willem (2007) Parsimonious data-oriented parsing. *Proceedings EMNLP 2007*, pp. 551–60.

Zwart, Jan-Wouter (2006) Uncharted territory? Towards a non-cartographic account of Germanic syntax. Paper presented at the 21st Comparative Germanic Syntax Workshop, Santa Cruz. Forthcoming in the *Proceedings* to be published by Benjamins. (http://www.let.rug.nl/~zwart/docs/cgsw21proc.pdf)

Zwarts, Joost (2007) Number in Endo-Marakwet. In Mechthild Reh and Doris Payne (eds.), *Advances in Nilo-Saharan Linguistics: Proceedings of the 8th Nilo-Saharan Linguistics Colloquium* (University of Hamburg, 22–25 August 2001), pp. 281–94. Cologne: Köppe.

Zwicky, Arnold (1994) Dealing out meaning: fundamentals of grammatical constructions. *Berkeley Linguistics Society* 20: 611–25.

This page intentionally left blank

Language Index

Name Index

Subject Index