OXFORD LINGUISTICS

edited by Frederick J. Newmeyer and Laurel B. Preston

Measuring Grammatical Complexity

structure local evolution modeling global dependency description length parsing variation quantification hierarchy parameter creoles sign language processing trade-off universality

Measuring Grammatical Complexity

Measuring Grammatical Complexity

Edited by

FREDERICK J. NEWMEYER AND
LAUREL B. PRESTON

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

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

© editorial matter and organization Frederick J. Newmeyer and Laurel B. Preston 2014
© the chapters their several authors 2014

The moral rights of the authors have been asserted

First Edition published in 2014

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2014932091

ISBN 978–0–19–968530–1

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Contents

Acknowledgments

The editors would like to thank the Andrew W. Mellon Foundation for providing Professor Newmeyer with an Emeritus Fellowship during academic years 2011–2013. This fellowship funded the Workshop on Formal Linguistics and the Measurement of Grammatical Complexity, held in Seattle on March 23–24, 2012. Most of the chapters in this volume originated as talks presented at the workshop.

We are also indebted to Allison S. Germain, a graduate student in Linguistics at the University of Washington, for her excellent work in preparing the indexes to this volume.

Notes on the contributors

Theresa Biberauer is a Senior Research Associate and a Fellow of Churchill College at the University of Cambridge, and Associate Professor Extraordinary at Stellenbosch University. Her research interests are principally in theoretical and comparative (synchronic and diachronic) morphosyntax, with Germanic generally and Afrikaans in particular being areas of specific interest. Past work has focused on word-order variation, (null) subject phenomena, negation, information structure, and the larger question of the nature of parametric variation.

Damián Blasi is a Ph. D. candidate in the Research Group of Discrete Biomathematics at the Max Planck Institute for Mathematics in the Sciences and the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology. His research involves the use of statistical methods and applied mathematics for the analysis of large-scale cross-linguistic patterns in the interface between linguistics and the behavioral sciences.

Cristiano Chesi is Researcher at the Center for Neurocognition and Theoretical Syntax at the Institute for Advanced Study (IUSS) in Pavia, Italy. He teaches computational linguistics at IUSS and at the University of Siena and collaborates with the Microsoft Language Development Center in Lisbon. He is interested in computational complexity and has developed a 'top-down, left-right' formalization of Minimalist grammar.

Peter W. Culicover is Humanities Distinguished Professor in the Department of Linguistics at The Ohio State University and a Fellow of the American Academy of Arts and Sciences and the Linguistic Society of America. His research has been concerned with explaining why grammars are the way they are. Culicover has worked in recent years on grammar and complexity, the theory of constructions ('syntactic nuts'), the history of the core constructions of English, and ellipsis.

Cecily Jill Duffield received her Ph. D. in 2013 from the Department of Linguistics at the University of Colorado at Boulder. She is currently conducting research in theoretical models of language production with an emphasis on the formation of grammatical dependencies. Her research interests include construction grammar and the interface of syntax and semantics in language production.

David Gil obtained his Ph. D. from UCLA in 1982 and has held teaching positions at Tel Aviv University, Haifa University, National University of Singapore, and Universiti Kebangsaan Malaysia. Since 1998 he has been at the Max Planck Institute for

Evolutionary Anthropology in Leipzig, for which he established and has been running the Jakarta Field Station. David's main interests are linguistic theory, linguistic typology, and field work, with a focus on the languages of Indonesia.

John A. Hawkins is Professor of Linguistics at the University of California Davis and Emeritus Professor of English and Applied Linguistics at Cambridge University. He has held previous positions at the University of Southern California, the Max-Planck-Institute for Psycholinguistics in Nijmegen, and the University of Essex. He has broad interests in the language sciences and has published widely on language typology and universals, efficiency and complexity in language, psycholinguistics, the Germanic language family, and language change.

Anders Holmberg is Professor of Theoretical Linguistics at Newcastle University, UK. His research interests are primarily syntactic theory and comparative syntax, with forays into descriptive syntax and typology. Much of his work is on the Scandinavian languages and Finnish. Recent projects include null subjects, word order universals, theoretical approaches to linguistic variation, and the syntax of polar questions and their answers.

Ray Jackendoff is Seth Merrin Professor of Philosophy and Co-Director of the Center for Cognitive Studies at Tufts University. His interests span syntax, semantics, social cognition, music cognition, and consciousness. He is the only theoretical linguist to have been awarded the Jean Nicod Prize in cognitive philosophy and the David Rumelhart Prize in cognitive science. He is the author of 14 books, of which the most recent is *A User's Guide to Thought and Meaning* (2012).

Lisa Matthewson is a Professor of Linguistics at the University of British Columbia. The goal of her research is to uncover what understudied languages can tell us about variation and universality in natural language semantics. She is currently conducting fieldwork on, and analysis of, St'át'imcets (Salish), Gitksan (Tsimshianic), and Niuean (Polynesian).

Lise Menn is Professor of Linguistics at the University of Colorado. She works on psycholinguistic models of aphasic morphosyntax in conversation and elicited narrative, and of phonology in early first-language acquisition. She is the author of *Psycholinguistics: Introduction and Applications* and the co-editor of *Agrammatic Aphasia*: *A Cross-Linguistic Narrative Sourcebook* and of *Methods for Studying Language Production*.

Steven Moran is a post-doctoral researcher in the Department of General Linguistics at the University of Zurich and a member of the research unit Quantitative Language Comparison located in the Deutsche Sprachatlas group at the University of Marburg. His research interests include language documentation and description in West Africa (Dogon and Gur families), phonological typology, and the development and application of quantitative methods for historical-comparative analysis.

Andrea Moro is Professor of General Linguistics and Director of the Research Center for Neurocognition and Theoretical Syntax at the Institute for Advanced study at IUSS in Italy. He has been a Fulbright student and visiting scientist at MIT and Harvard and has obtained a Diplôme d'études supérieures at the University of Geneva. Among his books are *The Raising of Predicates* (1997), *Dynamic Antisymmetry* (2000), *The Boundaries of Babel* (2008), and *The Equilibrium of Human Syntax* (2013).

Frederick J. Newmeyer is Professor Emeritus at the University of Washington and Adjunct Professor at the University of British Columbia and Simon Fraser University. He is Past President of the Linguistic Society of America and a Fellow of the American Association for the Advancement of Science. His primary linguistic interests are syntax and the history of the field.

Laurel B. Preston is a Ph. D. candidate in the Department of Linguistics at the University of Washington, Seattle. Her research interests include temporal semantics, with particular attention to the temporal interpretation of nouns, and adult second language acquisition. Her current work investigates poverty-of-the-stimulus problems in second language acquisition.

Ljiljana Progovac is Professor and Director of Linguistics, Wayne State University. Her research interests include syntax, language evolution, and Slavic syntax. In addition to a number of articles and chapters, she has published three books on syntax, including *Negative and Positive Polarity* (1994) and *The Syntax of Nonsententials* (2006). Since 2006, she has given 22 presentations on the topic of the evolution of syntax, in eight countries, and published 17 papers on the topic.

Ian Roberts has been Professor of Linguistics at the University of Cambridge, England since 2000. He has to his credit six monographs and two textbooks, and several collections of articles. He was joint Editor of *Journal of Linguistics* 1994–8, President of GLOW 1993–2001, and President of the Societas Linguistica Europaea 2012–3. His research interests are syntactic theory and comparative and historical syntax. He is currently working on a project funded by the European Research Council called 'Rethinking Comparative Syntax'.

Daniel Ross is a Ph. D. student at the University of Illinois at Urbana-Champaign. His fundamental interest in linguistics is how language works as a system, which has lead to research projects in syntax and semantics, typology, and language acquisition, in addition to his studying a variety of languages.

Michelle Sheehan is a Research Associate at the University of Cambridge specialising in comparative syntax with a particular interest in the Romance languages. She has worked on null arguments, control, word order variation, extraposition, clausal-nominal parallels and case/alignment. She is co-editor of *Theoretical Approaches to Disharmonic Word Order* (2013) and co-author of *Parametric Variation: Null Subjects in Minimalist Theory* (2010) and *The Philosophy of Universal Grammar* (2013).

Kaius Sinnemäki is currently working as a research fellow at the Helsinki Collegium for Advanced Studies, University of Helsinki. He wrote his dissertation *Language Universals and Linguistic Complexity* in 2011 and co-edited the volume *Language Complexity* (2008) with Matti Miestamo and Fred Karlsson. His current research interests include the cross-linguistic distribution of grammatical complexity, its sociocultural variation, and its relation to processing complexity.

Andreas Trotzke is a postdoctoral researcher in the Department of Linguistics, University of Konstanz. He received his Ph. D. in linguistics from the University of Freiburg in 2010. His research interests include left-peripheral syntax, the syntax-pragmatics interface, sentence processing, and the recursivity of syntactic embedding.

Eva Wittenberg is a doctoral candidate in linguistics at Potsdam University and a visiting fellow at Tufts and Harvard universities. Her research interests are conceptual representations and linguistic structure, which she studies using a variety of experimental techniques.

Jan-Wouter Zwart received his Ph. D. from Groningen in 1993. He is professor of theoretical linguistics at the University of Groningen and director of that university's Graduate School for the Humanities. His research and teaching centers around syntactic theory, linguistic typology, and Germanic linguistics. He has been a visiting scientist at the Massachusetts Institute of Technology and a visiting professor at the University of Geneva. His recent publications include the monograph *The Syntax of Dutch* (2011).

List of figures

List of tables

List of abbreviations

1/2/3	1st/2nd/3rd person
ABS	absolutive
ACC	accusative
AOR	aorist
APPL	applicative
DAT	dative
DEF	definitive
DEIC	deictic
DEON	deontic
DET	determiner
DIR	direct
DIST	distal
EMPH	emphatic
EPIS	epistemic
ERG	ergative
EVID	evidential
EVID.COP	evidential copula
EXIS	assertion of existence
F	feminine
FOC/TNS	particle of focus/tense
GM	generalizing modality
IOBJ	indirect object
INF	infinitive
IPFV	imperfective
INTR	intransitive
INTS	intensifier
MED	medial (verb form)
NMLZ	nominalizer
NOM	nominative
OBJ	object
PL	plural

POSS	possessive
PFV	perfective
PRED	predicative
PRS	present
PSB	possibility
PST	past
RED	redirective applicative
SBJ	subject
SG	singular
STAT	stative
SUBJ	subject
TAM	tense, aspect, modality
TEMP	temporal
TOP	topic
TR	transitivizer
TRNS	transitive

1

Introduction

FREDERICK J. NEWMEYER AND LAUREL B. PRESTON

1.1 Overview

The topic of the relative complexity of languages, along with its measurement, has begun to attract a great deal of interest. The past few years have seen the appearance of edited volumes (e.g. Miestamo *et al.* 2008; Sampson *et al.* 2009; Baerman *et al.* 2014), a single-authored book (Culicover 2013), and literally dozens of journal articles and anthology chapters devoted to questions around complexity and complexity measurement. The present volume contributes to the growing body of literature, but approaches the problem of the measurement of grammatical complexity from a unique direction. Most complexity researchers to date have backgrounds in either functional linguistics or sociolinguistics, or have applied issues around complexity to support some particular framework for syntactic analysis. What we offer here are discussions of the measurement of grammatical complexity from a variety of approaches to formal linguistics and from studies of the brain. We feel that the interested reader will find that the chapters here complement nicely the already-published material on the topic. We devote the first half of this introduction to a historical review of ideas about the relative complexity of languages. The second half gives an overview of the thirteen chapters that follow.

1.2 On the idea that 'all languages are equally complex'

Popular wisdom has always dictated that languages can vary from each other dramatically in terms of their degree of relative complexity.[1] To provide an illustrative example, the 1956 edition of the *Guinness Book of World Records* 'identified' the world's 'most primitive language' (Sampson 2009a: 227). The choice was the Australian language Arunta (now generally referred to as Aranda), in which 'words are indeterminate in meaning and form.' Despite everyday opinion, however, by the late

[1] Much of the material in this section is treated in greater detail in Joseph and Newmeyer (2012). See also Hurford (2012: ch. 5) for remarks that are largely in accordance with those of this chapter.

twentieth century a consensus had arisen among professional linguists that all languages are equally complex. The first quote that we have found that explicitly puts forward the idea that all languages are equally complex is from an article published in 1954 by Rulon Wells:

> Again, one can isolate the complexity of a language in phonemics, in morphophonemics, in tactics, etc.; but these isolable properties may hang together in such a way that the total complexity of a language is approximately the same for all languages. (Wells 1954: 104)

Just a year after the Wells article, the idea of equal complexity had found its way into the *Encyclopedia Britannica*. The following quote is from the entry on 'Language' written by George L. Trager:

> All languages of today are equally complex and equally adequate to express all the facets of the speakers' culture, and all can be expanded and modified as needed. There are no 'primitive' languages, but all languages seem to be equally old and equally developed. (Trager 1955: 698)

It is now standard for introductions to linguistics to make that same point. Consider an example from a popular introductory text:

> There are no 'primitive' languages—all languages are equally complex and equally capable of expressing any idea in the universe. (Fromkin and Rodman 1983: 16)

The hypothesis of equal complexity has by now appeared in a treatise on language change (McMahon 1994: 324), a published lecture on the relationship between language and thought (Bickerton 1995: 67), a book on Indo-European language and culture (Fortson 1997: 4), and a popular book about recent genetic discoveries (Ridley 1999: 94).

Three independent currents converged to make the hypothesis practically inevitable by the 1950s.[2] The first is a by-product of classical humanism: since all human groups are in a fundamental sense 'equal', their languages must be 'equal' too. The pioneers of early twentieth-century linguistics, including Boas, Sapir, and Bloomfield, focused on the *lack of correlation* between complexity and culture, rather than on equal complexity per se. However, given the generally accepted (by the 1940s) position that all languages could be analyzed by the same methods, it was a short (although not logically necessary) step to the conclusion that they do not differ in terms of overall complexity. Indeed, by the end of the twentieth century it came to be accepted by virtually all linguists on humanistic grounds that the very idea of a complexity ranking for languages was racist in its theoretical foundations and abhorrent in its application to real world practice. As David Gil put it tersely:

[2] The arguments based on humanism and language processing are discussed briefly in Maddieson (2007).

[S]ome people seem to think that if one language were shown to be more complex than another, then it would follow that the latter language is in some sense inferior, which in turn would entail that the speakers of that language are inferior, and from here we're only one short step to ethnic cleansing. (Gil 2001a: 326)

The second current derives from considerations of language processing. The idea is that in order to keep languages useable, complexity in one part of the grammar is necessarily 'balanced out' by simplicity in another part of the grammar. This is a very old idea. In fact, it is explicit in the following remark by Whitney in his *Life and Growth of Language*, first published in 1875:

[T]he means of formal expression are of the utmost variety; they are not to be sought in one department of a language only, but in all; they are scattered through the whole vocabulary, as well as concentrated in the grammatical apparatus. Deficiency in one department may be compensated, or more than compensated, by provision of resources in another. (Whitney 1897 [1875]: 222)

The idea of complexity trade-offs was hinted at or explicitly endorsed throughout the early twentieth century. Boas for example remarked:

Since, however, ideas must be expressed by being reduced to a number of related ideas, the kinds of relation become important elements in articulate speech; and it follows that all languages must contain formal elements, and that their number must be the greater, the fewer the elementary phonetic groups that define special ideas. In a language which commands a very large, fixed vocabulary, the number of formal elements may become quite small. (Boas 1963 [1911]: 27)

And André Martinet emphasized over and over again that languages represent an uneasy balance resulting from the conflicting needs of their users and from the consequences of having to deal with purely physical asymmetries in the vocal tract:

At every stage, the structure of language is nothing but the unstable balance between the needs of communication, which require more numerous and more specific units [. . .] and man's inertia, which favors less numerous, less specific, and more frequently occurring units. It is the interplay of these two main factors that constitutes the essentials of linguistic economy. (Martinet 1962: 139)

Hockett explicitly gave a language use-based explanation for his belief that languages do not differ significantly in terms of complexity:

Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology [word structure] and syntax [sentence structure], is about the same as any other. This is not surprising, since all languages have about equally complex jobs to do, and what is not done morphologically [. . .] has to be done syntactically [. . .]. Fox [. . .] , with a more complex morphology than English, thus ought to have a somewhat simpler syntax; and this is the case. (Hockett 1958: 180)

Since all languages have 'the same job to do' (whatever that might mean concretely), it follows, according to Hockett, that complexity in one area must be balanced by simplicity in another.

Examples of complexity tradeoffs are not hard to find. One of the best known is at least a century old. Rich case marking tends to correlate with flexible word order (see Siewierska 1998 for the relevant statistics). Consider the history of those Indo-European languages which have lost much case marking but have developed more rigid order. Or take a phonological example: Uncontroversially, we think, complex syllable structure correlates with low tonal complexity (Matisoff 1973). Another example that has been cited to support the idea of complexity trade-offs is Chinese and typologically similar languages, which have a simple (isolating) morphosyntax and individual morphemes that are multiply ambiguous. Such languages tend to have both constructions situated at the intersection between the lexicon and productive syntax like classifiers, reduplication, compounding, and verb serialization (see Riddle 2008 for Hmong, Mandarin, and Thai) and complex rules of inference and complex rules interfacing form and meaning (see Bisang 2009 for Khmer, Thai, and late Archaic Chinese). Complexity trade-offs continue to be proposed to this day, their existence bolstering the idea of all languages manifesting the same degree of overall complexity.

The third piece of motivation for the idea of equal complexity of all languages is derived from the (putative) demands of universal grammar. Chomsky has never explicitly asserted that all languages are equally complex, though he has come close. In the course of a discussion of the 'cost' of language-particular lexical peculiarities like having a special preposition to mark grammatical roles (where 'cost' seems equivalent to 'complexity'), Chomsky's interviewers interrupted his narrative with a typologically-oriented question: "All languages ought to be equally costly, in this sense?" Chomsky replied: "Yes, they ought to be" (Chomsky 2004a: 165–166). Some of his closest supporters have been more explicit. The first citation below is from a semi-technical exposition of Chomsky's ideas whose opening two sentences are "Why is Chomsky important? He has shown that there is really only one human language: that the immense complexity of the innumerable languages we hear around us must be variations on a single theme" (Smith 1999: 1). The second citation is from a technical work that contains a glowing Foreword by Chomsky:

> Although there are innumerable languages in the world, it is striking that they are all equally complex (or simple) and that a child learns whatever language it is exposed to. (Smith 1999: 168)

> Similarly, if we assume biologically determined guidance [in language acquisition], we need to assume that languages do not vary in complexity. (Moro 2008: 112)

The first concerted challenges to the idea of equal complexity came from the direction of sociolinguistics, rather than grammatical theory. Since the 1980s sociolinguists have suggested that the degree of complexity of a language might be in part due to factors such as its degree of contact or lack of contact with other languages, its relative prestige vis-à-vis its neighbors, its overall number of speakers, its role as a lingua franca, and other sociopolitical considerations (for one of the first publications along these lines, see Trudgill (1983).[3] While no sociolinguists, to our knowledge, have proposed precise complexity metrics, it would appear that linguists focusing on language and culture had never ceased to be a receptive audience for the view that languages might differ in complexity.

As we have seen, grammatical theorists have in general been much more skeptical than sociolinguists over whether languages can differ from each other in their degree of complexity. Two issues regarding complexity have engendered a firestorm of debate within the world of (primarily generative) grammatical theory: the relative complexity of creole languages and the question of whether some languages might lack recursive rules. We examine them in turn.

A tradition going back over a century holds that there is something 'special' about creoles that exempts them from generalizations applying to other natural languages. Traditionally, it was considered uncontroversial that creoles are simpler than non-creoles (see, for example, Schuchardt 1980); the controversies involved making precise the particular sociohistorical circumstances giving rise to the simplification. Derek Bickerton adapted the creoles-are-simpler hypothesis into generative grammar by positing that such languages manifest (or come close to manifesting) the most unmarked parameter settings allowed by universal grammar (Bickerton 1984; see also Roberts 1999). The hypothesis that the lexicon is the locus of parametric differences among languages allowed for creoles to be analyzed as simpler than non-creoles, while at the same time maintaining the hypothesis that, lexical differences aside, all languages manifest the same degree of complexity. However, there has been a determined resistance within the generative community to what is sometimes called 'creole exceptionalism' (see Muysken 1988; DeGraff 2001). It has to be said that this debate has at times passed the limits of what is normally considered acceptable academic interchange, with veiled accusations of racism and worse. Be that as it may, it is our impression that most grammatical theorists, regardless of their orientation, would now agree that in some significant way creoles are 'simpler' than non-creoles.

Turning to recursivity, the standard position among linguistic theorists has long been that all languages have subordination (Fromkin and Rodman 1983: 123). Indeed, such an idea has been transformed into virtually axiomatic status. Hauser *et al.* (2002: 1569), putting forward the position of the Minimalist Program, "hypothesize that [the

[3] Trudgill was not the first to link the complexity of a language's grammar to external sociopolitical factors. For example, Jakobson (1971 [1929]: 82) had written decades before that 'It is not rare to observe that the tendency to simplify the phonological system grows in proportion to the range of use of a dialect, with the greatest heterogeneity among speakers of the generalized language'.

narrow faculty of language] only includes recursion and is the only uniquely human component of the faculty of language." It follows from such a hypothesis that a language lacking recursion is excluded from the realm of possibility for theoretical reasons. If such a language were found to exist, then it would not be far-fetched to describe it as 'simpler' than other languages. Daniel Everett claims to have found such a language, namely Pirahã, spoken in Brazilian Amazonia (see Everett 2005). Reinforcing his claim that Pirahã is less complex than other languages, Everett has also argued that this language lacks quantifiers, numbers, color terms, and much more. Partly because of the sheer audacity of the various claims and partly because the non-recursivity of Pirahã challenges directly what is now taken to be a defining characteristic of Chomsky's view of language, Everett's work has received attention in the non-scholarly press, as is attested by the review of Everett (2008) in the *New Yorker* (Colapinto 2007). Resistance to Everett's claims within the world of theoretical linguistics has been fierce (e.g. Nevins *et al.* 2009). Several of the papers in this volume discuss Pirahã and its implications for the measurement of relative complexity.

A number of generative linguists have exploited the idea that one parameter setting might be more marked than another, as a means of characterizing the differential complexity of one grammar vis-à-vis another. Some proposals involving complexity-inducing marked settings have treated preposition-stranding in English and a few other Germanic languages (van Riemsdijk 1978; Hornstein and Weinberg 1981), the inconsistent head-complement orderings in Chinese (Huang 1982a; Travis 1989), and unexpected (i.e. typologically rare) orderings of nouns, determiners, and numerals in a variety of languages (Cinque 1996). In a pre-parametric version of generative syntax, Joseph Emonds (1980) had hypothesized that verb-initial languages are rarer than verb-medial languages because their derivation is 'more complex', as it involves a marked movement rule not required for the latter group of languages. Baker (2001) reinterpreted Emonds's analysis in terms of marked lexical parameters. And Newmeyer (2011) has pointed out that every version of generative syntax has posited syntactic-like rules that apply in the 'periphery' or in the mapping from syntax to phonology and are hence exempt from the constraints that might force 'core grammar' or the 'narrow syntactic component' to manifest equal degrees of complexity in every language.

At the same time, research in the past few decades has cast doubt on the idea that complexity in one area of the grammar is necessarily counterbalanced by simplicity in another. For example, according to David Gil (2007, 2008, 2009), Riau Indonesian has (almost) no word-internal morphological structure, distinct syntactic categories, or construction specific rules of semantic interpretation. Gil insists that as a result of Riau's bare-bones structures, the typical sentence of this language is *vague*, not *ambiguous*, and hence Riau does not have more complex rules of semantic interpretation to compensate for its simple morphosyntax. Along the same lines, Ian

Maddieson (1984) has pointed to cases where we do not find trade-offs in phonology: Languages with large consonant inventories tend also to have large vowel inventories; few manner contrasts for stops and fricatives are not compensated for by more place contrasts; and languages with simpler segmental inventories tend to have less elaborate suprasegmental properties.

All of the above has contributed to an atmosphere where linguists of all stripes are increasingly willing to entertain the idea that one language might indeed be simpler or more complex than another.

The crucial question, of course, is how to actually measure and compare degrees of linguistic complexity. There have long been proposals for measuring the complexity of individual grammatical components. One thinks of Sapir's formulas for morphological typology (Sapir 1921 and its revision in Greenberg 1960). And Hawkins (2004), building on a tradition deriving from Miller and Chomsky (1963), has addressed and quantified morphosyntactic complexity from a processing point of view. The programmatic sketches for measuring complexity have ranged from what might be called 'grammar-based' to those that might be called 'user-based'. The former focus on elements of grammars per se and count the amount of structural elaboration, irregularity, and so on (for an example, see McWhorter 2007). The latter approach complexity in terms of the degree of difficulty for the user, whether the first-language acquirer, the second-language acquirer, or the adult user. For example, Trudgill (2011) suggests that the most promising method for measuring complexity would be to attempt to determine if some languages are more difficult for second-language learners than others, abstracting away from their degree of similarity. Very few proposals, however, factor in what might be called 'interpretive complexity', that is, the relative difficulty from one language to another for a hearer to assign a semantic and pragmatic interpretation to an utterance. Both grammar-based and user-based approaches are discussed in this volume, as is the question of interpretive complexity.

1.3 Chapter synopses

Given this historical and theoretical background, we now turn to a preview of the chapters.

Hawkins's chapter 2 sets the stage for the rest of the volume by making the case for the actual and potential contributions of formal linguistics to the study of grammatical complexity. He offers a number of examples where formalization based on syntactic primitives such as constituent structure enables a precise characterization of surface structure syntactic phenomena, thereby enabling precision of measurement. Once measured, different structures within a language can be ranked, as can be done for comparable syntactic structures cross-linguistically. Complexity rankings, in turn, can help to explain preferences, again both within and across languages. Hawkins advocates bringing this degree of precision of characterization to all levels of grammar, including morphology, phonology, and semantics. He argues that unless we investigate complexity across multiple levels, it will not be possible to specify the 'overall' complexity of a grammar. Hawkins also demonstrates how formalization contributes to the measurement of efficiency, a concept that he argues is more inclusive than complexity. Drawing on previous work, Hawkins shows how the additional factors considered under the combined umbrella of efficiency and complexity contribute to better predictions in areas of observed performance such as processing delays or preferences among alternate structures.

Although Hawkins argues that ranking languages by overall complexity is beyond our current tools, in chapter 3, Gil engages in just that task. He makes the case that sign languages and creoles are in fact on the simpler end of a continuum of overall grammatical complexity in the world's languages. Gil begins by evaluating the relative complexity of signs and creoles with respect to two particular morphosyntactic phenomena, specifically, the flagging of core arguments, and the expression of tense, aspect and modality (TAM). Considering a sample of 32 sign languages and 76 creoles, Gil shows that whereas core argument marking and TAM expressions are widespread in the world's languages, the former are mostly restricted or absent and the latter are mostly optional in signs and creoles, thus making these languages relatively simple with respect to these two phenomena. He argues that relative simplicity in both core argument marking and TAM marking can be unified as a manifestation of weakly grammaticalized predication, where predication is defined as a composite of thematic role assignment and headedness (Gil 2012b). Gil suggests that younger languages, signs and creoles being two examples, tend to have weaker predication and are thus simpler overall than languages with strongly grammaticalized predication. However, it is possible for older languages also to have weak predication, as is the case with Riau Indonesian.

The next two chapters argue that languages can internally manifest different degrees of complexity, where simpler structures co-exist with the more complex. The chapters then go on to discuss the implications of this idea for the evolution of language. Jackendoff and Wittenberg (chapter 4) present a hierarchy of syntactic complexity, beginning with a one-word grammar and developing models at levels of successively increasing complexity, up to a recursive phrase grammar. With further elaborations, this thought experiment could characterize fully complex languages. Since the authors take 'grammar' to be a mapping between sound and meaning, they formalize the semantic/pragmatic component in addition to the syntactic component of the grammars on the hierarchy. Throughout the hierarchy of grammars, there is a trade-off between syntactic complexity and semantic/pragmatic complexity. At the simpler syntactic levels, there is more reliance on pragmatics for successful communication, and there is correspondingly more reliance on syntax and less on semantics/pragmatics as the syntactic component becomes more complex. The authors briefly consider a number of natural language phenomena including, among others, home sign, pidgins and creoles, and late second language acquisition, and situate these phenomena on the proposed hierarchy. They note that fully complex languages make use of various levels of the hierarchy, which suggests that their layered approach to grammatical complexity is on the right track.

Progovac in chapter 5 also argues that grammars make use of structures of differing morphosyntactic complexity, even to express the same or similar meanings. Working within the minimalist framework, she defends the claims that root small clauses are syntactically simpler than Tense Phrases and that intransitive absolutive clauses are simpler than transitive clauses. These results suggest the possibility that syntax evolved slowly, rather than suddenly, with complex phenomena arising from the foundation of less complex structures. This gradual 'tinkering' has resulted in a layered grammar, where the increasingly complex constructions build on, rather than eliminate, their simpler predecessors. Progovac argues that it is not the case that all languages are equally complex, nor do languages necessarily compensate for simplicity in one domain with complexity in another. Progovac also reviews neurological evidence for differential syntactic complexity and, as did Hawkins, she invites more collaboration between formal approaches to complexity and related fields, including neuroscience and genetics.

Also working in a minimalist framework, Biberauer, Holmberg, Roberts, and Sheehan (chapter 6) consider the formal syntactic complexity of grammatical systems from the point of view of parametric theory. Under the assumptions that parameters are part of innately specified Universal Grammar and are set independently, parametric theory would predict that all grammars are equally complex. The authors challenge these assumptions. They propose that parameters are not genetically encoded, rather, they emerge in the course of acquisition. Parameters form hierarchies, with the least marked option at the top of the hierarchy; as a learner moves 'down' the hierarchy in response to input, the grammatical structures are characterized by more and more variation. Thus the 'lower' options are syntactically more complex than the higher ones. Given this perspective on parametric variation, the authors propose two different ways to quantify the overall complexity of a grammar. As a test case, they calculate the complexity of five different natural languages according to both metrics, which leads to interesting results for the question of whether or not grammars differ in complexity.

Trotzke and Zwart in chapter 7 present a minimalist approach to grammatical complexity, while noting some underlying similarities between minimalism and other approaches, such as Construction Grammar. They point out that in recent minimalist work, representational complexity is reduced in favor of a phase-based, derivational model. For example, where earlier versions of the minimalist approach would place topic and focus in the syntactic representation, Trotzke and Zwart argue based on German examples that narrow syntax does not encode information structure. Grammatical complexity arises from cyclical interactions of subderivations, each of which is very simple. The authors stress that the question of grammatical complexity must be asked separately for narrow syntax and for grammar as a whole, including the interfaces and their interactions. They propose that these separate questions receive separate answers.

Culicover's chapter 8 adopts the framework of Construction Grammar rather than the Minimalist Program. However, he shares the assumption that grammars prefer simplicity to complexity. If this is so, then how does complexity arise and why do complex phenomena persist, especially when less complex alternatives co-exist in the grammar? To answer these questions, he considers complexity from both formal and processing points of view. Formally, grammatical complexity increases with the number of idiosyncrasies and irregularities in a language. Using English infinitival relatives as an illustration, Culicover argues that the pressure to reduce formal complexity in a general domain can create idiosyncrasy in a sub-domain. From a processing point of view, complexity increases with the amount of resources recruited to process a sentence. Using Continental West German verb clusters as an example, Culicover proposes two dimensions of processing complexity: a scope bias and a dependency bias. He then shows how different word orders within the verb clusters vary in complexity. The more complex orders persist despite the presence of the less complex orders because different orders reflect different complexity biases. Finally, Culicover presents a set of computer simulations that illustrate how social factors such as differences in social network topology can act to either spread, sustain, or eliminate complex phenomena in a grammar. For instance, tightly-knit social networks that are resistant to change can overcome a grammatical bias against complexity, thereby leading to the retention of complex phenomena for a long time. In summary, Culicover argues for the explanatory power of his approach to grammatical complexity for questions of language change and language variation.

In chapter 9, Sinnemäki sets out a general method for measuring grammatical complexity, advocating a grammatical description-based approach that measures the formal complexity of subsystems of grammars. He adopts a definition of complexity as the length of the shortest description of a grammar, and then applies this metric in an investigation of the trade-off hypothesis, that is, the hypothesis that complexity in one area of the grammar is counterbalanced by simplicity in another. His test case is core argument marking, raising the question of whether grammars that use case marking tend to have free word order. Sinnemäki conducts three different experiments on a stratified sample of fifty languages, with each experiment utilizing a different description of grammar. Two of the experiments describe grammars in terms of 'resources' (Dahl 2004, 2008), that is, in terms of inventories of units and constructions. The third experiment describes grammars in terms of 'regulations', that is, constraints and requirements. The results offer support for the trade-off hypothesis, but only when grammars are described in terms of resources, not regulations.

Ross's chapter 10 argues against Sinnemäki and others who advocate measuring the complexity of subsystems of grammars. He presents evidence that grammatical complexity must be measured with respect to an entire grammar. Measuring a core subset of properties might be attractive from a practical standpoint, but may also yield an inaccurate result. Adopting a description length-based metric of complexity, Ross argues that peripheral phenomena can contribute to overall grammatical complexity. As a case in point, he gives a detailed discussion of *try and* pseudocoordination in English. Ross proposes that this phenomenon (though not necessarily types of pseudocoordination in other languages) requires the positing of an additional principle, the Inflectional Parallelism Requirement. All other things being equal, and holding the grammatical theory constant across the descriptions of the comparison languages, English will be more complex than other languages as a result of having this principle.

In another test of complexity trade-offs, Moran and Blasi (chapter 11) investigate the relative complexities of phonological systems crosslinguistically. In agreement with Sinnemäki, they consider but reject a user-based approach. Instead they focus on so-called absolute measures of phonological complexity such as segment inventory size. Moran and Blasi design experiments to overcome some methodological concerns observed in the existing literature. First, they compile a dataset of 1200 languages, which is much larger than that used in many previous studies. Second, some of their experiments make use of the minimal set of distinctive features instead of segments as the basis of quantification. Using these two design innovations, they calculate a number of different distributions of variables and correlations between variables, both linguistic and non-linguistic. They argue that distinctive features provide a better basis than segments for complexity rankings. They find significant but small correlations between variables such as consonants and vowels, mean word length and segment inventories, and latitude and syllable structure. Many of the significant correlations disappear when the authors restrict their attention to individual family stocks. Moran and Blasi conclude that there are no universally significant generalizations that can be made about the relative complexity of different subsystems of phonological systems of languages.

Matthewson's chapter 12 investigates semantic complexity, a relatively understudied topic. She begins by reviewing evidence of cross-linguistic semantic variation. For example, some languages have generalized quantifiers and others lack them. But does variation of this sort necessarily imply a difference in complexity? In order to address this question, Matthewson proposes possible definitions of and metrics for semantic complexity, including: formal complexity (the description length of formal representations); paradigm complexity (the number of distinctions encoded in a certain area of semantic space); and expressive complexity (whether or not the same total semantic space is covered). Her comparison of English and St'át'imcets on these metrics across such areas as quantifiers, determiners, and evidentials, among others, reveals an interesting distribution of relative complexity. Matthewson concludes that languages can indeed differ in overall semantic complexity.

The final two chapters are neurolinguistically oriented. Chesi and Moro in chapter 13 return us to the question of how formal linguistics can characterize grammatical complexity in general, and then investigate the correlation between formal complexity as defined by the Chomsky Hierarchy and brain activity. They argue that the Chomsky Hierarchy facilitates measures of processing complexity via the automata that model the various levels of grammar on the hierarchy. Complexity can be precisely quantified as a function of time (the number of state traversals required to process hierarchical embeddings) and space (the amount of memory needed to keep track of dependencies). But do these measures correspond well to neurological evidence with respect to complexity signatures in the brain? The authors present evidence that the correspondence is not direct. For instance, cross-serial dependencies are not necessarily harder to process than nested dependencies, despite the fact that nested dependencies fall lower on the Chomsky Hierarchy. As a caveat, Chesi and Moro point out that it is difficult to disentangle the roles of memory and hierarchical embedding when measuring brain activity. On the other hand, experiments have demonstrated that specific regions of the brain are more active when hierarchical embedding gets deeper or long-distance dependencies get longer. Chesi and Moro conclude by showing how the complexity findings and the neurological evidence can be modeled in terms of automata.

The volume concludes with chapter 14 by Menn and Duffield in which they argue that formal complexity measures can be judged successful only if they can correctly predict measures of neurological activity. Formal complexity on its own abstracts away from a number of factors that affect actual user difficulty, including among other things the differences between comprehension and production, and between skilled and novice users. The authors point out that we cannot assume that a theory-dependent measure of formal complexity will correspond to brain activity measures, further substantiating the evidence presented in the previous chapter. Menn and Duffield suggest that neurolinguistics needs to supply formal linguistics with a list of factors beyond formal structural considerations that affect complexity, such as paradigmatic complexity, frequency, and co-occurrence probabilities. The authors offer experimental evidence that these factors are in fact reflected in processing measures. Working towards a performance-compatible model of grammar, the authors introduce us to their model of speech production, called MISCHA (Model Integrating Sequential, Categorial, and Hierarchical Activation). MISCHA elaborates on previous models of on-line sentence processing, taking into account factors such as frequency and predictability, whose importance is attested in the evidence presented in this chapter.

These chapters demonstrate that formal linguistics has much to contribute to the study of grammatical complexity. Formal approaches, by their very nature, lend themselves to precise specification, which is clearly necessary as a prerequisite to any imaginable complexity measurement. However, a conclusion that can be drawn from the chapters to follow is the importance of validating formal proposals with respect to evidence from processing and brain activity. As noted, the formal proposals and the psychological and neurological evidence do not always match up. We see the next necessary step in better understanding the nature of grammatical complexity to be a greatly increased collaboration between formal linguists and researchers from other disciplines, in particular those devoted to the study of cognitive and neurological processes and to language in use.

2

Major contributions from formal linguistics to the complexity debate

JOHN A. HAWKINS

2.1 Introduction

In this chapter I lay out some of the more general contributions made by formal linguistics to discussions of grammatical complexity. I also discuss possible contributions that have not yet been exploited and that I believe should be further developed. In addition I shall mention some contributions that I see as having been less successful in defining complexity and in deriving correct predictions from it for grammars and language usage. The overall thrust of the chapter is to draw attention to the benefits of better integrating formal linguistic insights with those from neighboring branches of the language sciences when trying to define and understand complexity. In doing so, we can then make better predictions for grammars and their properties, as well as for the manner in which languages are processed in real time and learned. Specifically, I shall illustrate:

- how formal linguistics can give us a precise characterization of comparable syntactic structures cross-linguistically and of their relative complexity (§2.2);
- how it can define different levels of complexity in parsing phenomena such as garden paths in a way that psycholinguists have yet to exploit (§2.3);
- how it contributes key clarifications and insights to discussions of 'overall complexity' that are missing elsewhere and that need to be made forcefully (§2.4);
- how it enables us to define efficiency as well as complexity and how it makes possible metrics for measuring both (§2.5).

These four points will all focus on non-derivational surface structure descriptions of the structures in question. I argue furthermore:

- that rule-based metrics of complexity have been less successful than surface structure-based metrics in large part because of the lack of agreement between different models and their rule systems, from Minimalism (Chomsky 1995) at the one end to Simpler Syntax (Culicover and Jackendoff 2005) at the other (§2.6); and

- that formal linguistic insights need to be better integrated with psycholinguistic and computational insights and theories, to the mutual benefit of each (§2.7).

2.2 Structural complexity: the relative complexity of comparable structures

A formal characterization of comparable structures within and across languages can provide a complexity ranking for them. Much of my own work on syntactic variation has shown how these rankings can be used to account for preferred selections among alternative structures in performance, and for preferred grammatical conventions across languages (see Hawkins 1994, 2004, 2014). Key properties of the structures as described within formal linguistics are integrated in this approach with considerations of on-line processing as described in psycholinguistics and computational linguistics.

I discuss two examples here: alternative word orders, and alternative relative clauses. When languages permit choices between different word orders, the alternatives can be described in terms of the relative sizes of their 'Phrasal Combination Domains' (Hawkins 2004: 107). In describing these alternatives I assume the existence of syntactic primitives such as heads of phrases, projections from heads, phrase structure trees, etc., and I assume also a rapid and efficient on-line processor producing and parsing syntactic representations based on these primitives. Alternative relative clauses can be linked to the relative sizes of their 'Filler-Gap Domains' (Hawkins 2004: 175) and to the presence or absence of structural features like resumptive pronouns that make processing easier. Sections 2.1 and 2.2 present these examples, respectively.

2.2.1 *Alternative word orders*

Some orderings reduce the number of words that are needed to recognize a mother phrase M and its immediate constituent daughters (ICs), making Phrasal Combination Domains (PCDs) smaller and phrase structure processing faster. I define a PCD as follows:

(1) Phrasal Combination Domain (PCD)
The PCD for a mother node M and its I(mmediate) C(onstituent)s consists of the smallest string of terminal elements (plus all M-dominated non-terminals over the terminals) on the basis of which the processor can construct M and its ICs.

Compare (2a) and (2b):

(2) a. The man [$_{\mathrm{VP}}$ looked [$_{\mathrm{PP}_1}$ for his son] [$_{\mathrm{PP}_2}$ in the dark and quite derelict building]]
1 2 3 4 5

b. The man [$_{\mathrm{VP}}$ looked [$_{\mathrm{PP}_2}$ in the dark and quite derelict building] [$_{\mathrm{PP}_1}$ for his son]]
1 2 3 4 5 6 7 8 9

The mother phrase VP consists of three items, V, PP1, PP2 that can be recognized on the basis of five words in (2a), compared with nine in (2b), assuming that (head) categories such as P immediately project to mother nodes such as PP, enabling the parser to construct and recognize them on-line. The principle of Early Immediate Constituents (Hawkins 1994) predicted a selection preference in direct proportion to the relative sizes of these PCDs:

(3) Early Immediate Constituents (EIC) [Hawkins 1994: 69–83]
The human processor prefers linear orders that minimize PCDs (by maximizing their IC-to-word ratios), in proportion to the minimization difference between competing orders.

EIC amounts to a preference for short before long phrases in head-initial structures like those of English, as seen in (2a). These orders will have higher 'IC-to-word ratios', i.e. they permit more ICs to be recognized on the basis of fewer words in the terminal string. The IC-to-word ratio for the VP in (2a) is 3/5 or 60% (5 words required for the recognition of 3 ICs). The comparable ratio for (2b) is 3/9 or 33% (9 words required for the same 3 ICs). For head-final languages like Japanese and Korean I predict a mirror-image long before short preference in performance, since the distances between their (rightmost) head categories are then kept just as minimal as they are in the short before long English structure (2a). (See Hawkins 1994: 80–81, 137–157 and Hawkins 2004: 108–110 for discussion and supporting evidence from these languages, and Hawkins 2014: ch.5.3 for more recent support.)

Structures like (2) were selected from an English written corpus on the basis of a permutation test (Hawkins 2000, 2004: 111–117): the two PPs had to be permutable with truth-conditional equivalence (i.e. the speaker had a choice). Only 15% (58/394) of these English sequences had long before short. Among those with at least a one-word weight difference (i.e. excluding 71 with equal weight), 82% had short before long, and there was a gradual reduction in the long before short orders, the bigger the weight difference (PPs = shorter PP, PPL = longer PP):

(4)

n = 323	PPL > PPs by 1 word	by 2–4	by 5–6	by 7+
[V PPs PPL]	60% (58)	86% (108)	94% (31)	99% (68)
[V PPL PPs]	40% (38)	14% (17)	6% (2)	1% (1)

This pattern of relative frequencies in (4) is what EIC predicts: the degree of preference for short before long PPs (i.e. with a more minimal PCD compared with the reverse long before short) is in proportion to the degree of difference between their respective PCDs, as measured by numbers of words and IC-to-word ratios. In a structure like (2) where PP2 is four words longer than PP1, the positioning of the shorter PP1 before the longer PP2 is found 86% of the time. When the difference between PPs is only one word (e.g. in the man looked [$_{PP1}$ for his son] [$_{PP2}$ in the dark building]) the preference for short before long is much less, at 60%. When the difference is greater than four words, the short before long preference is 94% or more. More generally, the more minimal these relative PCDs are, and the higher the corresponding IC-to-word ratios for one phrasal ordering versus the other, the more frequent the ordering is predicted to be.

This weight-based pattern is not the only one that I found in the corpus data of Hawkins (2000). I also explored semantic dependencies between verbs and prepositions (e.g. combinations like count on your father in which the meaning assigned to count depends on that following PP), whose processing requires a 'Lexical Domain' (cf. also Hawkins 2004: 117) sufficient for assigning the relevant properties listed in the mental lexicon in real time. In addition I explored pragmatic distinctions involving Givenness and Newness of information. I quantified the impact of all these factors on relative ordering. Syntactic weight (i.e. short before long orders with the most minimal PCDs) made the greatest number of correct predictions for these PPs and, when joined by semantic predictions for minimal lexical domains as well, achieved a success rate of 96%. Lexical domains for processing were minimal when count and on your father were adjacent and not separated by an intervening PP that did not figure in the lexical dependency relationship, such as in bad times in count in bad times on your father, which was dispreferred compared to count on your father in bad times. Pragmatic considerations were shown to be very weak compared with syntactic and semantic processing preferences defined in terms of PCDs and Lexical Domains. (See Wiechman and Lohmann 2013 for more recent data confirming that syntactic PCD sizes and semantic Lexical Domain sizes make the greatest number of correct predictions for post-verbal PP orderings and that pragmatic factors are weak. In contrast to Hawkins's 2000 these authors found that semantic factors were stronger than syntactic ones when the two were in competition.)

Because of the need to consider processing domains going beyond phrasal combination, for linear ordering as well as for other grammatical phenomena, I have subsequently generalized EIC into a Minimize Domains principle (see Hawkins 2004, 2013):

(5) Minimize Domains (MiD)

The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed. The degree of this preference is proportional to the number of relations whose domains can be minimized in competing sequences or structures, and to the extent of the minimization difference in each domain.

This formulation now defines a minimization and locality preference between all linguistic forms that are syntactically and semantically related to one another through grammatical operations or dependencies of any kind. It defines a gradient preference prediction for performance data such as (4) in proportion to the degree of syntactic weight difference between phrases, and in proportion to the number of syntactic and semantic relations that link them (there is a stronger adjacency preference between V and PP when they are phrasal sisters and also lexically dependent, compared with being just phrasal sisters alone).

MiD can also account for preference patterns across grammars, such as Greenberg's (1963) word order correlations. The numberings in (6) show the PCDs for a VP consisting of a V and a PP, from the first word of the domain to the last, in four possible ordering alternatives:

(6) a. $[_{VP}$ went $[_{PP}$ to the movies]]
      1 2
   b. [[the movies to $_{PP}$] went $_{VP}$]
      1 2
   c. $[_{VP}$ went [the movies to $_{PP}$]]
      1 2 3 4
   d. $[[_{PP}$ to the movies] went $_{VP}$]
      1 2 3 4

The adjacency of V and P in (6a) and (6b), i.e. with consistent head-initial or head-final ordering, guarantees the smallest possible string of words for recognition and construction of this VP and its two immediate constituents (ICs), V and PP. Two words suffice for their recognition and construction, i.e. these orders have the highest possible IC-to-word ratios of 2/2 = 100%. The inconsistent orders in (6c) and (6d) each require the processing of four words for recognition and construction of VP, resulting in a lower IC-to-word ratio of 2/4 = 50%. Matching this, cross-linguistic samples show that consistent head ordering of V and P is massively preferred over inconsistent ordering, as seen in (7) (which assumes a 2-word NP and a single-word V and P, as in (6), thereby assigning one more word to the phrasal node NP):

(7) a. $[_{VP}$ V $[_{PP}$ P NP]] = 161 (41%)
      IC-to-word: 2/2 = 100%
   b. [[NP P $_{PP}$] V $_{VP}$] = 204 (52%)
      IC-to-word: 2/2 = 100%
   c. $[_{VP}$ V [NP P $_{PP}$]] = 18 (5%)
      IC-to-word: 2/4 = 50%
   d. $[[_{PP}$ P NP] V $_{VP}$] = 6 (2%)
      IC-to-word: 2/4 = 50%
   MiD-preferred (7a)+(b) = 365/389 (94%) [Data from Dryer's 1992 sample]

The other Greenbergian correlations pattern similarly. For instance, the structure in (8) involves the head categories P and N within a PP and with N modified by a possessive phrase (e.g. with friends of my father in English). In languages with prepositions the head noun before possessive order is preferred within NP (see (8a)), in languages with postpositions the preference is for the reverse possessive before head noun (8b):

(8) a. $[_{PP}$ P $[_{NP}$ N Possp]] = 134 (40%) b. [[Possp N $_{NP}$] P $_{PP}$] = 177 (53%)
c. $[_{PP}$ P [Possp N $_{NP}$]] = 14 (4%) d. $[[_{NP}$ N Possp] P $_{PP}$] = 11 (3%)
MiD-preferred (8a) + (b) = 311/336 (93%) [Data from Hawkins 1983]

Many other grammatical ordering universals can be explained by MiD in this way (Hawkins 1994: ch.5, 2004: ch.5). It is significant also that Dryer (1992) points to certain systematic exceptions in Greenberg's correlations (i.e. in the kinds of word order preferences illustrated in (7) and (8)). There is less consistent head ordering when the category that modifies a head of phrase is a single-word item, e.g. an adjective modifying a noun (yellow book). Many otherwise head-initial languages have non-initial heads here (such as English) and many otherwise head-final languages have noun before adjective (e.g. Basque). But when the non-head is a branching phrasal category (e.g. an adjective phrase as in English books yellow with age) there are good correlations with the predominant head ordering. Why should this be?

The answer follows from Minimize Domains (5). When a head category like V has a phrasal sister, e.g. a PP in {V, PP}, then the distance from the higher head to the head of the sister will be very long when heads are inconsistently ordered and separated by a branching phrase, e.g. $[_{VP}$ V [NP P $_{PP}$]]. An intervening phrasal NP between V and P makes the PCD for the mother VP long and inefficient. But when heads are separated by a non-branching single word, then the difference between, say, $[_{VP}$ V $[_{NP}$ N Adj]] (read books yellow) and $[_{VP}$ V [Adj N $_{NP}$]] (read yellow books) is short, i.e. only one word. Hence the MiD preference for noun initiality (and for noun-finality in postpositional languages) is significantly less than when there are intervening branching phrases, and either less head ordering consistency or no consistency is predicted. When competing word orders in performance have PCDs that differ by just one word we have seen (in the first column of the double PP data in (4)) that both ordering options are productive, i.e. both short before long and long before short orders are frequent. Similarly when competing orders within and across grammars differ by just one word, as they do when a single-word adjective precedes or follows its head, both orders are productive as well.

The point here is that a formal characterization of these word orders in terms of phrasal groupings and distinctions like head versus non-head categories enables us to identify descriptive patterns in performance and grammars. And in combination with efficiency principles of a psycholinguistic nature like EIC (3) and MiD (5), we can then predict and explain many subtle distinctions in performance preferences and in grammatical universals which would be impossible in the absence of this formal characterization.

2.2.2 *Relative clause structures*

A similar point can be made with respect to the grammar and usage of relative clauses. For example, Keenan and Comrie (1977) proposed an Accessibility Hierarchy (AH) for describing patterns in the productivity of relative clause formation across languages and in the selection of different types of relative clauses. The original formulation of the AH has been reduced and changed slightly in subsequent publications by these authors. In the present context I shall focus on the version given in Comrie (1989), reproduced here as (9):

(9) Accessibility Hierarchy (AH): SU > DO >IO/OBL > GEN

Illustrative sentences from English for these four relativizable positions on the AH are given in (10):

(10) SU the student [who worked hard]
DO the book [which the professor wrote]
IO/OBL the course [which the professor wrote the book for]
GEN the professor [whose book the student read]

Keenan and Comrie presented evidence that relative clause formation may cease to apply across languages, or 'cut off' as they called it, in a systematic manner down the AH. Some languages like Malagasy permit relativization only on a SU, others only on SU & DO, yet others on SU & DO & IO/OBL, etc., as shown in (11):

(11) Rules of relative clause formation and their cut-offs within the clause
SU only: Malagasy, Maori
SU & DO only: Kinyarwanda, Indonesian
SU & DO & IO/OBL only: Basque, North Frisian, Catalan
SU & DO & IO/OBL & GEN: English, Hausa, Hebrew

These languages permit no relative clauses to be formed below the AH position shown. Keenan and Comrie also showed that the selection of different types of relative clauses, or 'strategies' as they called them, exhibited further patterns within the set of grammatical relative clauses across languages. For example, the distribution of 'gap' to 'resumptive pronoun' strategies followed the AH, i.e. in relative clauses corresponding to the book$_i$ which the professor wrote 0_i (with a gap as in English) versus the book$_i$ which the professor wrote it$_i$ (this resumptive pronoun being found in languages like Hebrew). Gaps favored the higher positions of the AH, resumptive pronouns the lower positions in Keenan and Comrie's data in languages that had both (see the precise formulation of this pattern below).

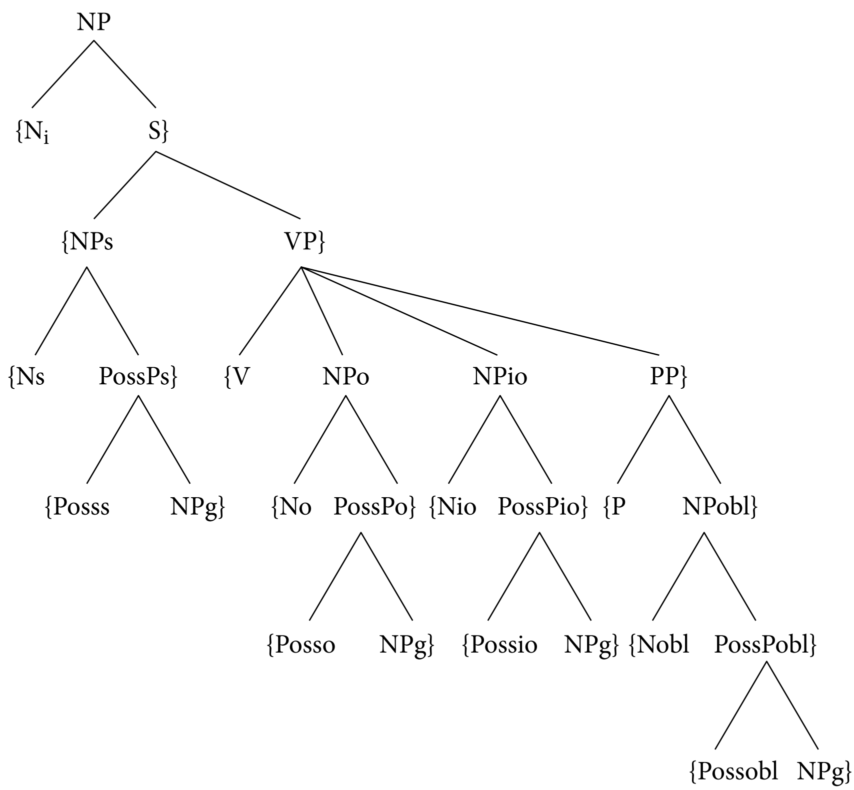
I have argued (Hawkins 1994: 37–42, 1999, 2004: 177–190) that Keenan and Comrie's data can be explained in terms of the increasing structural complexity of different relativizable positions down the AH, described in terms of expanding 'Filler-Gap Domains' (FGDs). An FGD is defined in (12) (taken from Hawkins 2004: 175):

(12) Filler-Gap Domain (FGD)
An FGD consists of the smallest set of terminal and nonterminal nodes dominated by the mother of a filler and on a connected path that must be accessed for gap identification and processing; for subcategorized gaps the path connects the filler to a co-indexed subcategorizor and includes, or is extended to include, any additional arguments of the subcategorizor on which the gap depends for its processing; for nonsubcategorized gaps the path connects the filler to the head category that constructs the mother node containing the co-indexed gap; all constituency relations and co-occurrence requirements holding between these nodes belong in the description of the FGD.

This complexity claim for relativization is not straightforward, given the differences between languages in terms of how their relative clauses are structured. But there is a clear general correlation between the number of dominating and co-occurring nodes required for the positions relativized on and their AH ranking. **Subjects** generally occupy the highest argument position in a clause and c-command other NP positions (generally asymmetrically). A **direct object** requires the co-occurrence of a (c-commanding) subject and often contracts a tight bond with the verb motivating a dominating VP in addition to S. FGDs for object gaps always contain the subcategorizor's subject, whereas FGDs for subject gaps need not contain objects. An **indirect object** generally requires the co-occurrence of both subject and object and in this context I assume that both are present in FGDs with indirect object gaps. Indirect objects are not distinguishable from obliques in terms of the complexity metric proposed here. **Oblique** NPs are often embedded within a PP, which gives them extra depth, but they differ from indirect objects in not generally requiring a direct object, though they do require a subject. A **genitive** NP is dominated by a possessive phrase within a higher dominating NP in a variety of structural positions, the aggregated complexity of which puts the genitive at the bottom of the hierarchy.

The configurational structure for different relativizations down the AH can be summarized in the single tree representation of (13), with unordered ICs for each phrase ('s' identifies constituents of a subject NP within the relative, 'o' constituents of an object NP, 'g' identifies the genitive NP within the relevant NPs, NPo, etc.):

(13)



In (14) I calculate the minimal FGD sizes for (gap) relativizations on each of the AH positions, given these assumptions (see Hawkins 2004: 177–190 for full details). The minimal FGD will include:

- the filler N_i (and NP);
- the co-indexed subcategorizor of N_i's gap (V_i, P_i or $Poss_i$, and their projected VP, PP, PossP);
- any overt arguments (NP and head N) that are required by the gap position (for instance, DO requires SU);
- (for genitives) the NP and/or PP dominating the subcategorizor $Poss_i$ and PossP plus the heads N and/or P that project to NP and PP;
- and the verb in the relative clause (and its dominating VP and S).

(14) Minimal FGDs for relativizations on:

SU = 5	{N_i, NP, V_i, VP, S}
DO = 7	{N_i, NP, V_i, VP, NPs, Ns, S} (requires SU)
IO = 9	{N_i, NP, V_i, VP, NPs, Ns, NPo, No, S} (requires DO & SU)
OBL = 9	{N_i, NP, P_i, PP, NPs, Ns, V, VP, S} (requires SU)

GEN-SU = 9	{N_i, NP, $Posss_i$, PossPs, NPs, Ns, V, VP, S}
GEN-DO = 11	{N_i, NP, $Posso_i$, PossPo, NPs, Ns, NPo, No, V, VP, S} (requires SU)
GEN-IO = 13	{N_i, NP, $Possio_i$, PossPio, NPs, Ns, NPo, No, NPio, Nio, V, VP, S} (requires DO & SU)
GEN-OBL = 13	{N_i, NP, $Possobl_i$, PossPobl, NPs, Ns, NPobl, Nobl, PP, P, V, VP, S} (requires SU)

The total number of nodes in the minimal FGD increases down (14). As these node counts increase there are more structural relations and co-occurrence requirements to compute, and more morphosyntactic and semantic dependencies. There will also be more terminal elements in larger domains, with more words to access and process. Node quantity is just one index of the relative complexity of an FGD, but it is a fundamental one with numerous correlating properties that involve additional processing operations.

This formal characterization of relative complexity can make sense of Keenan and Comrie's cut-off data given in (11): grammars cut off in proportion to the added complexity of their FGDs. I have also argued (Hawkins 2004) that a formal structural characterization in conjunction with the efficiency principle MiD (5) can explain the remarkable reverse distribution of gaps and resumptive pronouns in Keenan and Comrie's data: gaps prefer the easier, and resumptive pronouns the more complex FGDs, as shown in Table 2.1. More precisely, if a gap occurs on a low position in the AH, then gaps are found on all higher positions as well; if a resumptive pronoun is found on a high position, then resumptive pronouns are found on all lower positions that can be relativized at all (see Hawkins 2004: 186–190).

TABLE 2.1 Languages combining [−Case] gaps with [+Case] pronouns (Keenan-Comrie 1977)

	SU	DO	IO/OBL	GEN
Aoban	gap	pro	pro	pro
Arabic	gap	pro	pro	pro
Gilbertese	gap	pro	pro	pro
Kera	gap	pro	pro	pro
Chinese (Peking)	gap	gap/pro	pro	pro
Genoese	gap	gap/pro	pro	pro
Hebrew	gap	gap/pro	pro	pro
Persian	gap	gap/pro	pro	pro

Continued

Table 2.1 Continued

		SU	DO	IO/OBL	GEN
Tongan		gap	gap/pro	pro	pro
Fulani		gap	gap	pro	pro
Greek		gap	gap	pro	pro
Welsh		gap	gap	pro	pro
Zurich German		gap	gap	pro	pro
Toba Batak		gap	*	pro	pro
Hausa		gap	gap	gap/pro	pro
Shona		gap	gap	gap/pro	pro
Minang-Kabau		gap	*	*/pro	pro
Korean		gap	gap	gap	pro
Roviana		gap	gap	gap	pro
Turkish		gap	gap	gap	pro
Yoruba		gap	gap	0	pro
Malay		gap	gap	RP	pro
Javanese		gap	*	*	pro
Japanese		gap	gap	gap	gap/pro
Gaps	=	24 [100%]	17 [65%]	6 [25%]	1 [4%]
Pros	=	0 [0%]	9 [35%]	18 [75%]	24 [96%]

Key: gap = [−Case] strategy
pro = copy pronoun retained (as a subinstance of [+Case])
* = obligatory passivization to a higher position prior to relativization
0 = position does not exist as such
RP = relative pronoun plus gap (as a subinstance of [+Case])

[−Case] gap languages may employ a general subordination marker within the relative clause, no subordination marking, a participial verb form, or a fronted case-invariant relative pronoun. For Tongan, an ergative language, the top two positions of AH are Absolutive and Ergative respectively, not SU and DO, cf. Primus (1999).

The AH defines a ranking for relativizing on different positions *within* a clause. I have proposed similar hierarchies for relative clause gaps that are linked to head noun fillers *across* clause boundaries in Hawkins (1999, 2004: 192–201). One such is given in (15) with languages illustrating the relevant cut-offs shown in (16):

(15) A clause-embedding hierarchy for relative clause gaps

infinitival (VP) complement > finite (S) complement > S within a complex NP

(16) <u>Relative clause gap cut-offs across clause boundaries</u>

Infinitival (VP) complement:	Swedish, Japanese, English, French, Russian, German
Finite (S) complement:	Swedish, Japanese, English, French
S within complex NP:	Swedish, Japanese

For example, Standard German permits relativization on a gap position in an infinitival (VP) complement, as shown in (17), but not on a gap position in a finite (S) complement as in (18) (see Kvam 1983; Hawkins 1986):

(17) Wir diskutierten das Buch$_i$ [das$_i$ ich [$_{VP}$ [0$_i$ zu finden] versucht hatte]]/[das$_i$ ich versucht hatte [$_{VP}$ [0$_i$ zu finden]]] (German)
we discussed the book that I tried to find

(18) *der Tierpark$_i$ [den$_i$ ich vermute [$_S$ [dass alle Deutschen kennen 0$_i$]]] heisst…
the zoo which I imagine that all Germans know

Russian has conventionalized a similar distinction (see Comrie 1973):

(19) Vot ogurcy$_i$ [kotorye$_i$ ja obeščal [$_{VP}$ [prinesti 0$_i$]]] (Russian)
here are the cucumbers which I promised to-bring

(20) *Vot ogurcy$_i$ [kotorye$_i$ ja obeščal [$_S$ [čto prinesu 0$_i$]]]
here are the cucumbers which I promised that I'd bring

English allows relativization on a gap in a finite (S) complement, as in (21), but not on an S within a complex NP. Otherwise put, English obeys the Complex NP Constraint (Ross 1967), as illustrated in (22). Swedish does not. Relative clauses like (23), which corresponds to the ungrammatical (22), are grammatical in Swedish (Allwood 1982; Engdahl 1982).

(21) I was looking for a bone$_i$ [which$_i$ I know [$_S$ that the dog was gnawing on 0$_i$]]

(22) *I was looking for a bone$_i$ [which$_i$ I saw [$_{NP}$ a dog [$_S$ that was gnawing on 0$_i$]]]

(23) ett ben$_i$ [som$_i$ jag ser [$_{NP}$ en hund [$_S$ som gnager på 0$_i$]]] (Swedish)
a bone which I see a dog which is-gnawing on

These and many other complexity rankings predicted by MiD (5) are defined and illustrated in Hawkins (2004: ch.7). They require a formal characterization of the 'domains' for relative clause processing along the lines of (12) and (14). In other words, numerous details of a formal grammar involving the surface structures of relative clauses, which are either shared between different grammatical models or for which there are corresponding formal properties in different models, are vital for understanding linguistic usage and processing and grammatical patterns across languages.

2.3 Giving precision to parsing operations

The insights and methods of formal grammar can add precision to discussions of ease and difficulty in language processing, and can help make improved predictions for language performance under experimental conditions. For example, we can give a quantification based on formal grammatical criteria for the severity of what are generally called 'garden paths' in the parsing literature (see e.g. Frazier 1985), examples of which are given in (24) and (25):

(24) I believe the boy knows the answer

(25) The horse raced past the barn fell

In Hawkins (1994: 245–248) I linked a large number of such garden paths to considerations of structural complexity, and in Hawkins (2004) I discussed them in the context of the general theory of efficiency developed there (see section 5 below). I referred to garden paths as 'on-line misassignments' and I defined a quantitative metric for them (see Hawkins 2004: 49–61). A misassignment is severe and causes processing difficulty, I argued, in proportion to the amount of structural repair that needs to be implemented when correcting the initial parse. In (24), for example, the boy is initially misassigned to the matrix verb believe as a direct object, and is then subsequently assigned as a subject to knows within the subordinate clause. This is a less severe garden path than the notoriously difficult (25), in which much more of the structure is misanalyzed and then corrected only at the last word fell. The misanalyzed portions of sentences (24) and (25) are underlined.

Obviously there are other interacting factors that impact the severity or otherwise of a garden path in on-line processing, including real-world knowledge and frequency effects that favor certain on-line interpretations (see Macdonald *et al.* 1994 and the discussion below). These must all be taken into account in any multi-factor model that attempts to predict relative ease or difficulty in parsing (see Hawkins 2014: ch.9). But structural factors are a major contributor, and complexity formally defined has a key role to play in predicting and better understanding the experimental results involving on-line misassignments. In Hawkins (2004: 53–54) I defined the relative difficulty of an on-line misassignment as follows:

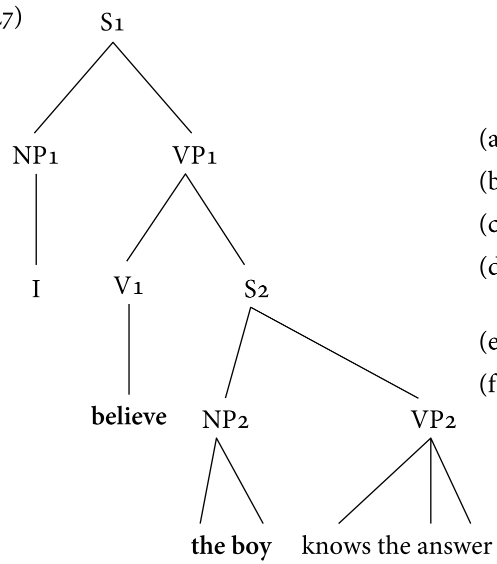
(26) Misassignment factors

a. the number of words and phrases that undergo some temporary misassignment of properties on-line;
b. the number of any additional dominating nodes that must be introduced into the syntactic tree when correcting the misassignments in (a);

c. the number of any mother-daughter attachments that are temporarily misassigned to the words and phrases in (a);
d. the number of any relations of combination or dependency that are temporarily misassigned to the words and phrases in (a);
e. the number of mother-daughter attachments that replace those misassigned in (c);
f. the number of relations of combination or dependency that replace those misassigned in (d).

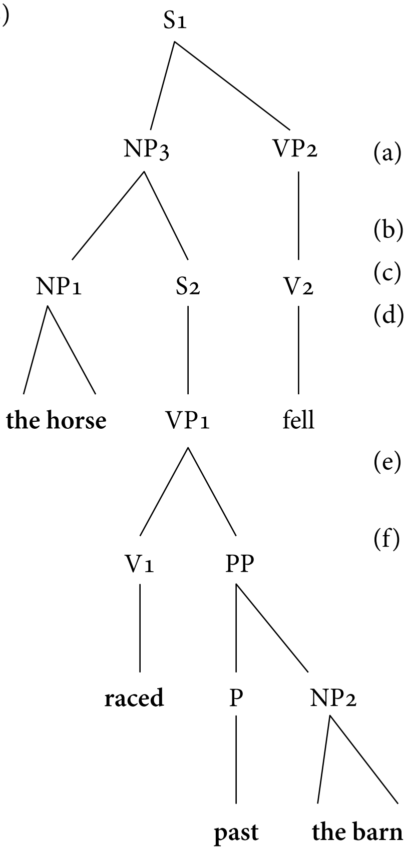
When these factors are added up and applied to different misassignment types it can be shown that (24) involves less restructuring than (25). Fewer words and phrases are misrecognized on-line in the former (see (26a)), fewer additional nodes must subsequently be introduced (26b), there are fewer misassigned relations of attachment (26c), combination and dependency (26d), and the number of such relations replacing those misassigned is less (26e) and (26f). This is shown in (27), which enumerates the misassignments in (24) and in (28), which enumerates the misassignments in (25). These misassigned entities, relations and properties have been quantified according to the criteria (a–f) in (26). (27) has a missassignment total of 11 such factors, whereas for (28) the total is much greater at 26. ([$_{VP1}$ NP2] means that NP2 is attached as an immediate daughter to VP1, etc.)

(27)



(a)	<u>the</u>, <u>boy</u>, NP2	[3]
(b)	S2	[1]
(c)	[$_{VP1}$ NP2]	[1]
(d)	NP2 = Accus of <u>believe</u> NP2 = θ-object of <u>believe</u>	[2]
(e)	[$_{S2}$ NP2], [$_{VP1}$ S2]	[2]
(f)	NP2 = Nomin of <u>knows</u> NP2 = θ-subject of <u>knows</u>	[2]
	TOTAL =	11

(28)



(a)	the, horse, raced, past, the barn, NP1, VP1, PP, NP2	[10]
(b)	NP3, S2	[2]
(c)	$[_{S1}$ NP1], $[_{S1}$ VP1]	[2]
(d)	VP1 = predication on NP1 VP1 = active VP NP1 = Nomin of raced NP1 = θ-subject of raced	[4]
(e)	$[_{NP3}$ NP1], $[_{S1}$ NP3], $[_{S2}$ VP1],$[_{NP3}$ S2]	[4]
(f)	VP1 = rel clause modifying horse VP1 = passive VP Nomin of raced= gap θ-subject of raced= gap	[4]
	TOTAL =	26

This kind of quantification of structural misassignments and their repair provides a metric for defining the relative complexity of different garden paths. It quantifies the important role of structural factors in the degrees of difficulty that have been recorded in processing experiments using these sentence types (see MacDonald *et al.* 1994 for an excellent review). This quantification can then be combined with quantifications based on the other relevant factors that MacDonald *et al.* discuss, such as frequency effects. For example, raced in the horse raced past the barn is more frequently used in the active than the passive, which sets up the garden path prior to the final verb fell in (25), leading to all the necessary structural repairs defined in (28). A verb like followed, on the other hand, in the horse followed past the barn is more frequently and more readily understood in the passive, which lessens the likelihood of a garden path in parsing. Linguistic context and real-world knowledge are also relevant here, and MacDonald *et al.* have shown how previous experimental findings can be better understood on the basis of a mix of quantitative and qualitative factors that interact in determining the severity or otherwise of garden paths. What is missing in their account is a precise quantification of structural factors and of structural complexity, going beyond intuitive assessments of complexity. The metric proposed here in terms of misassignments factors (26), which builds on standard formal linguistic analyses of the relevant properties, provides one and it can make our predictions for, and our understanding of, garden paths more precise as a result.

2.4 Defining 'overall complexity'

Formal grammars define and make explicit the mappings between sounds and meanings, via syntax. A structure that is simple in its surface syntax may actually be quite complex at another level of representation, or in its semantics. By describing these different levels of representation and the mappings to meaning for a given construction or structural type, formalization can serve as a corrective to claims in the literature about the overall simplicity of certain structures, or even of whole grammars, based on only partial descriptions of their properties. I summarize one illustrative case here that is of relevance to the complexity debate since it has figured prominently in discussions in the literature. The structure I will discuss has been used as a paradigm example of grammatical simplicity. What it shows, in actual fact, is that simplicity at one level of representation can create complexity at another, which makes it impossible to provide any clear assessment of simplicity or complexity overall.

This example highlights an important general point: we do not yet have a metric of overall complexity, either for grammars as a whole, or even for one and the same structure when all of its component properties are considered.

English transitive clauses with NP-V-NP have simple adjacency of NPs and V without explicit prepositions or nominal case marking and have been claimed to be like those found in creoles and in the world's simplest grammars (McWhorter 2001). However, the English NP-V-NP is highly ambiguous syntactically and semantically and must often be mapped onto rather complex thematic roles and argument-predicate relations at the level of representation that defines these argument-predicate structures. Most other languages do not permit the kind of complexity that English exhibits here, not even the closely related German (see Rohdenburg 1974; Hawkins 1986; Müller-Gotama 1994; König and Gast 2007). Compare the following sentences of English and German. English regularly permits non-agentive thematic roles in subject position, like Instrument, Location and Theme. German does not.

(29) a. *The professor* wrote an important book. [Agent]
b. *Der Professor* hat ein wichtiges Buch geschrieben.

(30) a. A few years ago *a penny* would buy two to three pins. [Instrument]
b. *Vor einigen Jahren kaufte *ein Pfennig* zwei bis drei Stecknadeln.

(31) a. *This tent* sleeps four. [Location]
b. **Dieses Zelt* schläft vier.

(32) a. *The book* sold 10,000 copies. [Theme]
b. **Das Buch* verkaufte 10,000 Exemplare.

Instead German translates these non-agentive subjects with less ambiguous and more semantically transparent prepositional phrases, as shown in the German counterparts (34) for English (33):

(33) a. *This advert* will sell us a lot of dog food.
b. *Money* can't buy everything.
c. *This statement* overlooks the fact that the situation has changed.
d. *This* loses us our best midfield player.
e. *The lake* prohibits motor boats.
f. *The Langdon (river)* could and frequently did drown people.
g. *My guitar* broke a string mid-song.
h. *The latest edition of the book* has added a chapter.
i. ... Philips, *who* was streaming blood, ...

(34) a. *Mit dieser Werbung* werden wir viel Hundefutter verkaufen.
b. *Mit Geld* kann man nicht alles kaufen.
c. *Mit dieser Aussage* übersieht der Autor, dass ...
d. *Damit* verlieren wir unseren besten Mittelfeldspieler.
e. *Auf dem See* sind Motorboote nicht zugelassen.
f. *Im Langdon* konnte man ertrinken, was übrigens häufig genug vorkam.
g. *An meiner Gitarre* riss mitten im Lied eine Seite.
h. *Zur letzten Ausgabe* des Buches wurde ein Kapitel angefügt.
i. ... Philips, *von dem* Blut herunter strömte, ...

These contrasts reveal the following. Transitive and ditransitive surface structures in English are syntactically and morphologically simple. The NP-V-(NP)-NP structure is minimal and contains fewer syntactic categories than its German counterpart with PPs. English lacks the case morphology of German, except residually on its pronouns. The rules that generate these English surface forms are also, arguably, more minimal than their German counterparts and generate a larger proportion of the outputs of English grammar (see Dahl 2004). Yet in any formalization of the mappings from surface forms to argument-predicate structures, and from these to semantic representations, the set of English mappings must be regarded, by anyone's metric, as more complex than the German set. There are more argument-predicate types to be linked to NP-V-(NP)-NP than to the corresponding transitive and ditransitive structures of German (containing NPs only and without PPs). This means more thematic role combinations for the same surface structure type, which adds 'length' and complexity to the grammar of English. It also makes processing more complex. The assignment of an Instrument to the subject NP in (30a) and of a Locative in (31a) requires access by the processor to the verbs in these sentences, to their lexical semantics and co-occurrence structure, and also possibly to the post-verbal NP. More disambiguation needs to be carried out by the English processor, and greater access is needed to more of the surface structure, since the processing domains for theta-role assignments in English are less minimal.

It is true, therefore, that the superficially simple SVO orders of English are found in inflectionally impoverished languages and in creoles. But any claim about the overall simplicity of such structures must take into account a concomitant increase in complexity elsewhere.

More generally, whereas complexity can be measured precisely within a given area, in terms of numbers of structural nodes and relations etc., we do not yet have a method of assessing the simplicity or complexity of a whole construction, let alone of a whole grammar (see Deutscher 2009 for this same point). Most of the examples of relative complexity illustrated in the present chapter have compared alternative versions of just one and the same structure at a single level of representation, essentially in its surface syntax. This, I have argued, is an achievement in and of itself. But we now see that other levels of representation must be taken into account as well if we are to classify a particular construction type, or a whole grammar, as simple or complex overall. And the problem is that we do not currently have a way of doing this and of reconciling simplicity and complexity at different levels of representation. The contributions to discussions of complexity that I have exemplified in previous sections have accordingly focused on single levels of representation, mainly surface syntax. What is clear is that if an overall metric is ever to be developed, formalization on all levels will be a necessary prerequisite for it. In the meantime claims about the overall simplicity of whole grammars, or even of whole constructions, should be made with caution.

2.5 Defining efficiency

I have argued (Hawkins 2004, 2009, 2014) that discussions of complexity need to be embedded within a theory of efficiency, and that it is efficiency that is the larger and more inclusive notion. Efficiency relates to one of the most basic functions of language, which is to communicate information from the speaker (S) to the hearer (H), and to the manner in which this function is performed. I propose the definition in (35):

(35) Communication is efficient when the message intended by S is delivered to H in rapid time and with the most minimal processing effort that can achieve this communicative goal.

and the following hypothesis:

(36) Acts of communication between S and H are generally optimally efficient; those that are not occur in proportion to their degree of efficiency.

Complexity metrics, by contrast, are defined on the grammar and structure of language itself. An important component of efficiency often involves structural and grammatical simplicity, as argued above in §2.2 and 2.3. But sometimes efficiency can actually result in greater complexity.

For example, if the speaker can refer to a given entity or event with a shorter form or structure, this is generally more efficient than using a longer and more complex one (see e.g. Ariel 1990). The pronoun he is formally simpler than a full and complex referential phrase like the professor who was teaching my linguistics class this morning. But in the event that speaker and hearer do not share the knowledge base that permits the simpler he to be successful in discourse, then the more complex expression is required. On these occasions it is not efficient for the speaker to use he, since this does not achieve the communicative goal, namely successful reference.

More generally, efficiency involves factors in addition to structural complexity that determine the speaker's selections of forms and structures, including:

- fine-tuning forms and structures (their degrees of explicitness, their presence or absence) to frequency of occurrence, accessibility and inference (see Ariel 1990; Rohdenburg 1996; Wasow *et al.* 2011);
- (faster) speed in delivering linguistic properties in on-line processing; and
- reduced misassignments (i.e. garden paths) in on-line processing.

A focus on efficiency rather than on complexity enables us to model more of these factors that ultimately determine the preferences of performance and of grammars. It also helps us to better understand certain trade-offs between competing complexities and efficiencies. Consider some examples. A complex morphology (case marking in OV languages, rich verb-agreement in verb-early languages, see Hawkins 2004: 246–250) may provide faster on-line access to argument structure in processing. Greater structural complexity is actually efficient for processing in this case. An inconsistent head ordering and word order (NRel in an OV language, which creates a complex Phrasal Combination Domain, see (1) and Hawkins 2004: 205–209) may avoid garden paths and computationally expensive reconstructions of the tree. And deletions of material (zero allomorphs, reductions in phrase structure through compounding, gaps linked to fillers, controlled empty argument positions, and so on, see Hawkins 2004: chs.4 and 6) can all be better understood when linked to factors such as frequency of use, accessibility and ease of recoverability. Such deletions make these structures efficient, for reasons of economy and 'Form Minimization' (see Hawkins 2004: 38–49).

Formal linguistics provides the tools for analyzing structural phenomena such as these that can be linked to efficiency factors and it makes it possible for us to define principles of efficiency that are precise enough to be quantified and tested. The misassignment factors for garden paths and their severity discussed in §2.3, for example, defined degrees of *in*efficiency in on-line processing: the more structure-building that takes place that must ultimately be repaired, the more inefficient the parse is. Similarly in Hawkins (2004: 51–58) I defined a corresponding set of predictions for avoiding what I call 'unassignments'. These are delays in on-line processing that occur when using one structural variant compared with another. For example, if instead of ordering the reciprocal anaphor each other after its antecedent (as in I introduced the musicians$_i$ who were visiting the city to each other$_i$) I were to reverse the order through Heavy NP Shift (I introduced to each other$_i$ the musicians$_i$ who were visiting the city), there would be a delay in assigning the co-referential index to the reciprocal and in understanding its reference. By positioning and parsing the musicians$_i$ . . . first, the reference of each other$_i$ can be assigned immediately as it is parsed and a 'look ahead' to a subsequent item in the parse string is avoided. This preference for assigning as many as possible of the (correct) linguistic properties that are ultimately to be assigned to words in a parse string, as early as possible, is captured by the principle of Maximize On-line Processing (MaOP) which is reproduced here (see Hawkins 2004: 51):

(37) Maximize On-line Processing (MaOP)

The human processor prefers to maximize the set of properties that are assignable to each item X as X is processed, thereby increasing O(n-line) P(roperty) to U(ltimate) P(roperty) ratios. The maximization difference between competing orders and structures will be a function of the number of properties that are unassigned or misassigned to X in a structure/sequence S, compared with the number in an alternative.

Any delay in on-line processing through unassignment, with or without an additional misassignment, reduces the OP-to-UP ratio of a sentence. This ratio is calculated according to the procedure set out in (38):

(38) On-line Property to Ultimate Property Ratio

The OP-to-UP ratio is calculated at each word X within a connected set of words {. . . X . . .} whose property assignments differ between two competing structures S and S'. The cumulative number of properties assignable at each X is divided by the total number of properties to be assigned in the connected set in the ultimate representation of each structure, and the result is expressed as a percentage (e.g. 4/20=20%, 8/20=40%, 10/20=50% for successive words, etc.). The higher these on-line percentages, the more efficient is structure S or S', since more properties are assigned earlier.

The basic point here is that formalization is a prerequisite, not just for a theory of complexity, but for a broader theory of efficiency as well. Formalization helps us better understand the structures that are chosen or, alternatively, avoided in performance, and conventionalized or not conventionalized across grammars.

2.6 Rule-based formal metrics

Formal linguistics has been less obviously successful in defining complexity and its predictions for performance and grammars when the metric in question is based on rules and the derivations they generate, rather than on the structures themselves (especially observable surface structures). One early failure was the Derivational Theory of Complexity (DTC) (Miller and Chomsky 1963; Fodor, Bever, and Garrett 1974). According to the DTC, more complex derivations for sentences are supposed to result in longer reaction times in experiments, etc. It soon became apparent that sometimes they do, but often they do not!

A major problem in testing the DTC, of course, is that different generative theories have from the outset employed very different rule systems. In the early days of generative syntax, Adj-N combinations like interesting book could either be derived by various transformations from structures corresponding to book which is interesting or base-generated by Phrase Structure rules (cf. Brown and Miller 1996). There is even greater theoretical and descriptive diversity now. Compare Simpler Syntax (Culicover and Jackendoff 2005) at the one extreme with Minimalism (Chomsky 1995) at the other. The former prefers simplicity in its surface syntactic representations (with some added complexity elsewhere), while the latter has simple principles of derivation, but at the cost of very long and complex derivations for the generation of individual sentences.

In general, I would argue that the formal objects that have been most useful in defining simplicity versus complexity, and in making precise and correct predictions for structural preferences in performance and for cross-linguistic patterns of conventionalization, are those that are closest to observable surface structure, and not to their hypothesized derivations. Though theories do differ in the way they represent surface structure (see again Brown and Miller 1996), there seems to be more agreement among them in the (relative) complexity rankings they define for surface structure representations, and better empirical support (e.g. in terms of non-terminal to terminal node ratios, IC-to-word ratios, etc., see Miller and Chomsky 1963; Frazier 1985; Hawkins 1994, 2004; and Gibson 1998, 2000) than in the often very different derivations they postulate. Complexity can be relativized to whatever level of abstraction (numbers of empty categories, vacuous branching etc.) each theory postulates.

Consider an illustrative example. Rule-based mechanisms can often get in the way of the right surface structure generalization that is needed in order to capture complexity differences and make the right predictions for cross-linguistic grammatical patterns. In Hawkins (1983: 183–189) I borrowed the X-bar rule schema proposed in Jackendoff (1977) and expanded on his insight that the simplest grammars are those in which identical categories at common bar levels serialize in the same direction. This enabled him to formulate cross-categorial rules of ordering using just one single rule, rather than having separate rules with distinct orderings for the constituents of each phrasal category. The result was a generally good set of predictions for Greenberg's word order correlations exemplified in §2.1 above, with similar orderings for similar complements and their heads across categories: V-NP (VO) co-occurs with P-NP (prepositional) ordering, NP-V (OV) with NP-P (postpositional) ordering, etc.

The problem is, as Jackendoff notes (pp.81–85), that sometimes these rules get in the way of the required surface characterization of complexity differences. For example, the distinction between single-word terminal categories like Adj, and phrases like AdjP, is not easy to capture given that AdjP → Adj in PS-rules. Therefore, one cannot introduce an Adj into a phrase-structure tree in different environments from its dominating AdjP, given this PS-rule. Yet, as we have seen in §2.1, single-word items are often differently and inconsistently ordered across languages compared with their more complex phrasal counterparts, even though both are dominated by one and the same phrase such as AdjP. Compare English yellow book in a predominantly head-initial grammar with the phrasal and consistently ordered book yellow with age. Not only do rules get in the way of capturing the best formulation here—in terms of surface categories and their complexity irrespective of their domination by a common phrasal type—but the explanation for this and other facts of ordering has, arguably, nothing to do with the rule mechanisms per se, and involves instead the ease or difficulty of producing and recognizing the relevant surface structures.

2.7 Interdisciplinary integration for descriptive and explanatory adequacy

I have argued in this chapter that formal linguistics has a major role to play in the complexity debate by:

- giving precision to the description of comparable structures within and across languages and their relative complexity (§2.2);
- giving a precise characterization of complexity in parsing, e.g. in garden paths (§2.3);
- contributing key clarifications to claims of overall complexity (§2.4); and
- enabling us to define metrics of efficiency, e.g. for speed of processing, correctness in on-line parsing, and predictability and form minimization (§2.5).

Psycholinguists and computational linguists ignore these insights at their peril. At the same time, grammarians need to recognize that grammars are shaped by considerations of efficiency and complexity in language usage and communication to a greater extent than has traditionally been recognized, as is stressed in Newmeyer (1998, 2005) and Hawkins (2004, 2014). Hawkins (2004, 2014) proposes the following hypothesis and supports it with extensive evidence from the performance of different languages and from cross-linguistic grammatical data and variation patterns:

(39) Performance-Grammar Correspondence Hypothesis (PGCH)

Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments.

Improved descriptive and explanatory adequacy can come from admitting these insights from performance into our grammars, just as our formalizations can improve the processing mechanisms and computations defined in these neighboring fields.

3

Sign languages, creoles, and the development of predication

DAVID GIL

3.1 Introduction

Sign languages and creoles, although different from each other in important ways, have, over the past few decades, followed a broadly similar trajectory from the periphery into the mainstream of linguistic research. In the past, both sign languages and creoles were considered, for various reasons, to constitute less than 'real' languages. One of the major factors underlying this prejudice is an assumption that these languages are simpler than most others, lacking the grammatical complexity and expressive power of full-fledged languages. More recently though, an increasing body of empirical work, coupled with the shedding of non-linguistic biases, has led to the realization that both sign languages and creoles are of generally comparable complexity to most other familiar languages. For example, American Sign Language and Jamaican Creole are every bit as much real languages—and hence as much within the scope of mainstream linguistic theory—as, say, English and Mohawk. Thus, we are now in a position to move forward and ask whether sign languages and creoles are indeed like other languages in all respects (other than those that define them as sign languages and creoles, and of course those which make each and every language unique), or whether perhaps sign languages and creoles might differ from other languages in systematic ways with respect to certain grammatical properties.

While much of grammar is determined by universal cognitive factors, be they specific to human language or of a more general conceptual nature, it has always been widely assumed that other important properties of language and languages are shaped by historical and social factors. In particular, recent work by McWhorter (2011), Trudgill (2011) and others has provided empirical support for the claim that certain configurations of historical and social factors have an effect on the complexity of grammar, the generalization being that contact among adult speakers is conducive to simplification while isolation tends to bring about complexification. Indeed, McWhorter (2001, 2005) cites creole languages as providing the clearest case of an extreme contact scenario resulting in radical grammatical simplification, with strong further support coming from the large-scale cross-linguistic comparison by Parkvall (2008). But why should contact bring about simplification? The obvious answer is that when speakers of different languages come into contact with each other and try to communicate, they are forced to create a new shared system, which by its nature is an incremental process which starts out simple and accrues additional complexity as time goes by. The most celebrated example of this process, which linguists were able to document from the very beginning, is that of Nicaraguan Sign Language, which was formed in the course of just a very few years by children—each with his or her original system of home sign—who were brought together in a single school. But Nicaraguan Sign Language is unique amongst sign languages only in that we were able to see it happening. Most or all other sign languages presumably had comparable beginnings in the relatively recent past; Woll, Sutton-Spence and Elton (2001) estimate that no contemporary sign language is more than 300 years old. Accordingly, several sign-language scholars have suggested that sign languages in general resemble creoles, and may owe some of their grammatical features to their being relatively young, newly-formed languages—see for example Fischer (1978), Meier (1984, 2002), Gee and Goodhart (1985), Kegl, Senghas and Coppola (1999), Aronoff, Meir and Sandler (2000, 2005), Newport and Supalla (2000), and Sandler and Lillo-Martin (2005: 503–510).

This chapter shows that as young languages, sign languages and creoles tend to be simpler than other kinds of languages with respect to two central features of morphosyntax, namely the flagging of core arguments and the expression of tense-aspect-mood (TAM) categories. Building on the notion of predication proposed in Gil (2012b), which relates these two features, this chapter then argues that sign languages and creoles represent a relatively simpler grammatical type in which predication is either absent or only weakly developed, a type that is also observable in the language of young infants and may be reconstructable for that of our hominin ancestors.

It should be emphasized, however, that the greater simplicity of sign languages and creoles with respect to these two morphosyntactic features and the notion of predication underlying them does not set them apart from all other languages; some older languages may resemble sign languages and creoles in this respect. Rather, the claim is that with respect to these properties, sign languages and creoles occupy a proper subset of the space of typological variation spanned by all of the world's languages.

This chapter is organized as follows. Section 3.2 shows that across the world's languages, usual or obligatory flagging of core arguments and obligatory expression of TAM in basic simple clauses are both widespread. Sections 3.3 and 3.4 demonstrate that for sign languages and creoles, restricted or absent core-argument flagging and optional expression of TAM are the norm, concluding that such languages tend to be less complex than their older counterparts with respect to these two features. Section 3.5 then argues that the simplicity of sign languages and creoles with respect to the flagging of core arguments and the expression of TAM categories reflects the absence of a strongly gramamticalized notion of predication, and goes on to suggest that such languages may provide a model for early stages in the acquisition of language by children and the evolution of language itself.

3.2 Core-argument flagging and tense-aspect-mood marking

We begin with definitions of the two features under examination, core-argument flagging and TAM marking, and their possible feature values.

3.2.1 *Core-argument flagging*

The first feature, core-argument flagging, examines the extent to which the two core arguments of simple basic transitive clauses are flagged for their thematic role or grammatical function. The term 'core argument' includes the thematic roles of agent and patient, broadly construed so as to encompass also the experiencer and stimulus of verbs of perception; however, it excludes the additional role present in ditransitive constructions, as well as various others such as instrumental, locative and the like, that are generally assumed to be more of a more oblique nature. Attention is limited to full NPs; pronominal arguments are ignored. The term 'flagging' covers any marking of thematic roles on the nominal phrases that bear them; its most common manifestations are adpositions, clitics and affixes, though less frequently it may also involve non-linear mutations such as consonantal or vocalic modifications, or changes in lexical tone. Excluded from this definition of flagging are the marking of thematic roles on the verb or elsewhere in the clause, or in intonation. Flagging is taken to be present if it is overtly expressed on at least one of the two core arguments.

The feature of core-argument flagging classifies languages in accordance with two values:

(1) *Core-argument flagging*
- (a) restricted or absent (simple)
- (b) usual or obligatory (complex)

The contrast between these two types is exemplified by English, which instantiates restricted or absent core-argument flagging, as opposed to Cuzco Quechua, which is characterized by usual or obligatory core-argument flagging:[1]

[1] The interlinear glosses in this chapter make use of the following abbreviations: ACC accusative; DIR direct; EVID evidential; F feminine; INF infinitive; PRED predicative; PST past; SG singular; SUBJ subject; TOP topic; TRNS transitive; 3 3rd person.

(2) *English*
The man saw the woman

(3) *Cuzco Quechua* (Paul Haggerty pc)

Qhariqa	warmitam	rikurqan
man:TOP	woman:ACC:DIR.EVID	see:PST:3SG.SUBJ

'The man saw the woman'

Whereas in English and other languages, such as Ju|'hoan, Mandarin Chinese and Riau Indonesian, core-argument flagging is completely absent from full NPs in basic clauses, other languages represent an intermediate state of affairs in which core-argument flagging is optional. Some languages, such as Yali (a Trans-New-Guinea language of the Papua province in Indonesia) exhibit differential subject marking, in which an ergative marker occurs optionally on the agent; typically, the presence of the ergative marker is subject to various discourse pragmatic factors. Other languages, such as Spanish, exhibit the mirror-image case of differential object marking, in which an accusative marker occurs optionally on the patient; typically, the presence of the accusative marker is limited to patients that are animate, definite, or of otherwise high semantic or pragmatic salience. For present purposes, languages with differential subject or object marking such as Yali and Spanish are considered as having restricted core-argument flagging, in that the presence of core-argument flagging is limited to specific semantic or discourse contexts; they are accordingly grouped together with languages such as English as having restricted or absent core-marking flagging, as per (1a). Such languages thus contrast with Cuzco Quechua and other languages, including Hungarian, Georgian and Tagalog, which exhibit usual or obligatory core-argument flagging.[2]

To lay the ground for the survey of sign languages and creoles in the next two sections, the relative frequency of these two types across the world's languages must be ascertained. No such survey has yet been undertaken; however, an approximation thereof is provided by Nichols and Bickel's (2005) study on 'Locus of Marking in the Clause', which focuses on the coding of direct objects. Nichols and Bickel define four major language types: head marking, dependent marking, double marking, and no marking. The four types are defined in terms of the grammatical encoding of the direct object in relation to the verb: (i) on the verb; (ii) on the direct object; (iii) on both; or (iv) on neither. These four types may be collapsed into a binary distinction between their types (i) and (iv), lacking direct object flagging and their types (ii) and

[2] While overt core-argument flagging will often serve to differentiate between the two core arguments, this function is not taken to be definitional of core-argument flagging. Thus in Russian, for many verbs, nominative and accusative case markings are identical, and therefore do not distinguish between core arguments. However, since these case markings are generally overt, Russian is classified as having usual or obligatory core-argument flagging.

(iii), exhibiting direct object flagging. Nichols and Bickel's study, published in Haspelmath, Dryer, Gil, and Comrie (2005) *World Atlas of Language Structures (WALS),* has since been augmented with additional data classified in essentially the same way. This augmented database, provided by Balthasar Bickel, shows 156 languages of their types (i) and (iv) versus 133 languages of their types (ii) and (iii). Nichols and Bickel's classification provides a rough approximation to the classification in (1). Specifically, all or almost all of their type (i) and (iv) languages are languages with restricted or absent core-marking flagging, thus approximating class (1a). Note that they characterize a language as having a zero-marked accusative or absolutive case as having patient flagging, provided that it contrasts with an overtly marked agent. On the other hand, a majority of their type (ii) and (iii) languages are languages with usual or obligatory core-argument flagging, thus approximating class (1b). The exceptions are languages with differential subject or object marking, which Nichols and Bickel include in types (ii) or (iii) but—as explained above—are classified here as having restricted or absent core-argument flagging. Nevertheless, in spite of the mismatches between the two classifications, Nichols and Bickel's study suggests that it is safe to conclude that both of the language types in (1) are cross-linguistically widespread.

In terms of complexity, restricted or absent core-argument flagging represents a less complex state of affairs than usual or obligatory core-argument flagging, by dint of the extra flagging morpheme or morphemes that are present. Admittedly, other factors may point in alternative directions. For example, differential subject or object marking could be argued to be more complex than both the consistent absence or the obligatory presence of core-argument flagging, because of the additional semantic and/or pragmatic rules that govern the distribution of the relevant markings. More generally, it is not immediately obvious whether and to what extent the presence or absence of core-argument flagging has any bearing on the overall complexity of a language. For example, while Cuzco Quechua distinguishes core arguments by flagging, English does so by means of word order, and Skou (a language of the Skou family of Papua New Guinea) by means of verbal affixation, both alternative strategies introducing complexity in other parts of the grammar; no clear metric is available for ranking these alternative strategies in terms of complexity. Nevertheless, in the absence of such a metric, it would seem reasonable to adopt the default position whereby, all else being equal, restricted or absent core-argument flagging involves lesser complexity than usual or obligatory core-argument flagging.

It is often suggested that there are compensatory mechanisms whose function is to balance simplicity in one domain of grammar with complexity in another. Perhaps the most celebrated proposal of this kind is that which relates the loss of case marking in Romance languages with the rise of fixed word order; see for example Lehmann (1985b: 312) and Kroch (2001: 701). A typological study by Siewierska (1998) based on a sample of 171 languages shows that there is indeed a negative correlation between core-argument flagging distinctions and fixed word order; however, the correlation is a relatively weak one, with several counterexamples in either direction: languages with both core-argument flagging distinctions and fixed word order, and languages with neither. The latter kind of language is exemplified by Riau Indonesian, with neither core-argument flagging nor fixed word order, and where, in sentences such as the following, the interpretation is unspecified with regard to the assignment of agent and patient thematic roles:

(4) *Riau Indonesian*
Cowok tengok cewek
man see woman
'The man saw the woman' / 'The woman saw the man'

Constructions such as (4), with no formal core-argument distinction whatsoever, either by flagging or by means of other morphosyntactic devices, have a good claim to be simpler than their counterparts in other languages, such as those in (2) and (3). As argued in Gil (2008), languages such as Riau Indonesian suggest that there may indeed be cross-linguistic differences with regard to complexity in the broader domain of thematic roles and their encoding. In particular, as argued in Gil (2008), there is good reason to believe that in at least some of the relevant cases, speakers are not compensating for the absence of overt thematic-role marking by making greater use of their pragmatic competence.

3.2.2 *Tense-aspect-mood marking*

The second feature, TAM marking, pertains to the expression of TAM categories, making a binary distinction between optional TAM marking languages, in which there are some basic declarative affirmative main clauses with no grammatical expression of any TAM categories, and obligatory TAM marking languages, in which all basic declarative affirmative main causes contain a grammatical expression of at least one of the three TAM categories. These two types may be represented as follows:

(5) *Tense-Aspect-Mood Marking*
(a) optional (simple)
(b) obligatory (complex)

The contrast between these two types can be seen in (2)–(4) above. In languages such as Riau Indonesian, TAM marking is optional, and clauses often lack any expression of these categories, as is the case in (4). In contrast, in languages such as English and Cuzco Quechua, TAM marking is obligatory, with clauses always containing expression of at least one of these categories, as indeed is the case in (2) and (3).

For present purposes, the grammatical marking of TAM may be either bound or free. Moreover, it may be either dedicated, expressing nothing but TAM concepts, or portmanteau, combining the expression of TAM with that of other concepts. For example, in many Philippine languages, TAM is combined with voice, by means of distinct voice affixes associated with different aspects; similarly, in many European languages, TAM is combined with agreement features such as person, number and gender, via subject-verb agreement paradigms which vary in accordance with tense and/or aspect. In some languages, clauses with no overt TAM marking are constrained in their range of possible meanings; for example, in Yoruba, such clauses are interpreted as expressing either present or past time, depending on the aktionsart of the verb; see for example Welmers (1973: 346–347). If one considers TAM marking as constituting a single unitary paradigm, one might analyze such cases as involving a 'zero morpheme' expressing some particular TAM value. However, the present study takes the alternative what-you-see-is-what-you-get approach of characterizing such languages as possessing optional TAM marking. In other languages, though, the expression of TAM concepts via a verbal agreement paradigm may give rise to cases where some values of the agreeing features are zero-marked while others are associated with overt phonological content. A rather extreme example of this is provided by English with its simple present overtly marked only in the 3rd person singular, all other values being zero-marked. Since in such cases there is generally ample motivation for positing a paradigm containing zero-marked elements, such languages are characterized as possessing obligatory TAM marking.

A study of TAM marking across the world's languages in accordance with the above criteria is presented in Gil (to appear) and discussed further in Gil (2012b). The study is based on a sample of 868 languages; of these, 377 are categorized as having optional TAM marking, while 491 are classified as having obligatory TAM marking. What this study shows, then, is that both types are widespread across the languages of the world.

In terms of complexity, it is rather straightforwardly the case that, with the presence of fewer overt morphemes, optional TAM marking represents a less complex state of affairs than obligatory TAM marking. Moreover, it is not at all clear how languages with optional TAM marking might go about compensating for their simplicity here by introducing complexity elsewhere. One might imagine that they would opt for more frequent use of words or phrases referring to various abstract properties associated with activities such as their timing or internal structure, though there is no evidence suggesting that this is the case. Alternatively, one might suggest that their speakers make more extensive use of pragmatic inferences to fill in the purportedly 'missing' material, but once again, there is no evidence for this, that is, no reason to believe that speakers of a language such as Riau Indonesian always worry about whether a clause unmarked for TAM has past, present or future reference any more than speakers of a language such as English invariably puzzle over past-tense-marked clauses, wondering whether they refer to a few hours back, a day ago, a week to a month ago, a couple of months to a couple of years ago, or the distant or legendary past, just because these distinctions happen to be grammaticalized in a language such as Yagua (Payne and Payne 1990: 386–388).

3.2.3 *Correlating the two features*

The extent to which a language flags core arguments is logically independent of the extent to which it marks TAM categories. Indeed, examples can be found of languages exhibiting each of the four possible combinations of feature values in (1) and (5): Riau Indonesian (1a)/(5a); Casiguran Dumagat (1b)/(5a) (Headland 1974); English (1a)/(5b); Cuzco Quechua (1b)/(5b). Still, there is reason to believe that they correlate. Using Nichols and Bickel's (2005) survey as a substitute for core-argument flagging, one may test for a correlation with TAM marking. As shown in Gil (2012b: 326), such a correlation holds (at a level of significance $\alpha < 0.001$).[3] As argued there, this correlation reflects a deeper correlation between two more general properties, the extent to which thematic role assignment and tense are grammaticalized; and in turn, this latter correlation represents the degree of grammaticalization of the notion of predication. This correlation is discussed further in section 3.5.1 below.

3.3 Sign languages

The previous decade has witnessed an upsurge in studies examining the ways in which sign languages may differ from spoken languages with regard to their grammatical properties. Some of the proposed differences appear rather idiosyncratic. For example, Zeshan *et al.* (2012) point to various features of sign language numeral systems that are unattested in spoken languages, such as the 'digital' numerals of Indian Sign Language, in which a complex numeral such as, say, 512, is expressed as 'five one two', with no mention of the hundreds or tens. Similarly, Zeshan (2004: 30–31) observes that content questions lacking an overt interrogative word are common in sign languages whereas Gil (in progress) shows that they are very rare in spoken languages. Other proposed differences, however, are more far-reaching, and amenable to various kinds of principled explanation.

Some important differences between sign languages and spoken ones are accounted for in terms of modality effects. Thus, many scholars have drawn attention to the greater iconicity of sign languages, attributing this to properties pertaining to the visual medium—see the discussion in Sandler and Lillo-Martin (2005: 493–503). Similarly, many researchers have noted that sign language is more hospitable to

[3] The correlation between core argument flagging and TAM marking is also observable within languages, via the distinction between contrasting clause types such as non-finite vs. finite, or non-sentential vs. other; see Progovac (2006, 2009, this volume: chapter 5) and Gil (2012b: 322–323).

simultaneity than its spoken variety, the obvious explanation, once again, lying in the possibilities made available by visual space that are absent from the auditory medium—see Sandler and Lillo-Martin (2005: 491–493).

Other differences involve properties of sign languages that can be explained in terms of their relative youth, namely properties with respect to which sign languages typically resemble creoles. Thus, for example, Aronoff, Meir, and Sandler (2000, 2005) observe that sign languages are like creoles in that they typically have little in the way of concatenative morphology, because affixation has not yet had time to develop. More generally, Gee and Goodhart (1985) note that sign languages exhibit many of the properties attributed by Bickerton (1981) to creole languages, among which are several relating to the features under consideration here: like many creoles, sign languages have no adpositions, no tense marking, and no finite/non-finite distinction.

With the abovementioned studies providing a suitable context, sign languages may now be examined with respect to core-argument flagging and TAM marking. A total of 32 sign languages were examined. Due to the relative paucity of available information on many of the world's sign languages, there was no point in devising any kind of sampling strategy; every sign language for which information was available was included in the survey. The results of the survey are presented in Table 3.1:[4]

Table 3.1 Core-argument flagging and TAM marking in sign languages

		TAM marking		
		optional	obligatory	total
core-argument flagging	restricted or absent	32	0	32
	usual or obligatory	0	0	0
	total	32	0	32

[4] The 32 sign languages and their sources are as follows: Adamorobe Sign Language, Nyst (2007); Al Sayyid Bedouin Sign Language, Wendy Sandler (pc); Alipur Sign Language, Sibaji Panda (pc); American Sign Language, Wendy Sandler (pc); Australian Sign language, Johnston and Schembri (2007); Ban Khor Sign Language, Nonaka (2007); British Sign Language, Nick Palfreyman (pc); Catalan Sign Language, Quer and GRIN (2008); Chinese Sign Language, Fischer and Gong (2010); Deutsche Gebärdensprache, Ulrike Zeshan (pc); Hong Kong Sign Language, Fischer and Gong (2010); Indian Sign Language, Ulrike Zeshan (pc); Indonesian Sign Language, Nick Palfreyman (pc); Israeli Sign Language, Wendy Sandler (pc); Italian Sign Language, Zucchi, Neidle, Geraci, Duffy and Cecchetto (2010); Jordanian Sign Language, Hendriks (2008); Kafr Qasm Sign Language, Wendy Sandler (pc); Kata Kolok, Marsaja (2008); Kenyan Sign Language, Sam Lutalo-Kiingi (pc); Korean Sign Language, Kang-Suk Byun (pc); Língua de Sinais Brasileira, de Quadros, Müller and Lillo-Martin (2010); Nicaraguan Sign Language, Kegl, Senghas and Coppola (1999); Nihon Shuwa (Japanese Sign Language), Keiko Sagara (pc); Plains Indian Sign Language, Bakker (2012); Polish Sign Language, Paweł Rutkowski (pc); Taiwanese Sign Language, Smith (1990); Tanzanian Sign Language, Sam Lutalo-Kiingi (pc); Türk Isaret Dili, Ulrike Zeshan (pc); Ugandan Sign Language, Lutalo-Kiingi (in preparation); Vlaamse Gebarentaal, De Weerdt and Vermeerbegen (2008); Yolngu Sign Language, Dany Adone (pc); Zambian Sign Language, Sam Lutalo-Kiingi (pc).

The data in Table 3.1 present a clear-cut picture. All 32 sign languages have restricted or absent core-argument flagging and all 32 languages have optional TAM marking. With regard to the former feature, the generalization could perhaps be strengthened. In the course of the survey I encountered not a single instance of a sign language construction containing any kind of core-argument flagging, suggesting that 'restricted or absent' might be replaced by a simpler 'absent'. However, since much of the data in the survey is based on either fragmentary written descriptions or brief interviews with language experts, it seemed preferable to err on the side of caution.[5] With respect to the latter feature, while it has been suggested that sign languages generally lack tense, many are renowned for elaborate systems of verbal aspect. However, in all of the languages under consideration, it seems possible to identify an unmarked base verbal form to which aspectual inflections may apply as optional elaborations.

With respect to both core-argument flagging and TAM marking, sign languages thus differ strikingly from their spoken counterparts. While in spoken languages, usual or obligatory core-argument flagging is a widespread option, in these sign languages at least it is unattested. And while in spoken languages, obligatory TAM marking is a common feature, in the sign languages of the present sample it is completely absent.

Nevertheless, sign languages do not stand completely apart from spoken languages with respect to these two features; rather, they occupy a proper subset of the typological space of spoken languages. Consider the following example from Kata Kolok (a village sign language from Bali, Indonesia):

(6) *Kata Kolok* (Connie de Vos pc)
MAN SEE WOMAN
man see woman
'The man saw the woman'

With its absence of any core-argument flagging and TAM marking, Kata Kolok sentence (6) resembles similar sentences in many spoken languages. The difference between sign and spoken languages is that no sign language has either usual or obligatory core-argument flagging or obligatory TAM marking, two features that are characteristic of many spoken languages.

[5] An interesting problem is posed by intonation. In a study of American Sign Language, Fischer (1975) claims that pauses may be used to mark deviations from a basic SVO word order, which would then entail that intonation would be the sole conveyor of information about thematic roles. Similar claims are occasionally made about spoken languages. However, every such claim that I have been able to examine has turned out to be false—see Gil (2006) for discussion. If, however, Fischer is right about American Sign Language, then this would present yet another potential difference between sign and spoken languages. But even if she is right, the possible function of intonation in American Sign Language would not affect the results of the survey as presented in Table 3.1, since intonation is excluded from the definition of flagging.

Why do sign languages differ in this way from their spoken counterparts? The answer, it is suggested here, is because sign languages are young languages. To reinforce this point we turn to an examination of young languages in the spoken medium.

3.4 Creoles

Whereas most sign language researchers today would seem to be open to the possibility that sign languages may differ in systematic ways from their spoken-language counterparts, the situation in creole studies is quite different, with the thesis of 'creole exceptionalism' forming the subject of lively controversy. On one side are those who, often for ideological reasons, reject the possibility that creole languages may differ in significant ways from other older languages—see for example DeGraff (2003, 2005), Ansaldo and Matthews (2007) and others. On the other side are scholars mustering an array of empirical evidence suggesting that creole languages do in fact differ in systematic ways from their non-creole counterparts—see Bakker, Daval-Markussen, Parkvall and Plag (2011) for a recent large-scale cross-linguistic study.

A central theme in the creole exceptionalism debate is the notion of 'complexity'. While Bakker *et al.* deal with grammatical features that have no bearing on the notion of complexity, most of the empirical evidence in support of creole exceptionalism has revolved around the idea that creole languages are simpler than their non-creole counterparts—see McWhorter (2001, 2005), Parkvall (2008) and others. The motivation for this generalization is obvious, grounded in the history of creole languages as emerging from processes of radical social and linguistic disruption, pidginization, and imperfect second-language acquisition, all of which provide favorable environments for grammatical simplification.

To see how creole languages fare with respect to core-argument flagging and TAM marking, we turn to the Michaelis, Maurer, Haspelmath, and Huber eds., (2013) *Atlas of Pidgin and Creole Language Structure (APiCS)* database, which provides information on a sample of 76 languages. While a large majority of the *APiCS* languages are most probably creoles, a smaller number are pidgins, and a few belong to other kinds of contact languages such as mixed languages. The editors avoid assigning labels such as 'creole' or 'pidgin' to the sample languages on the grounds that such labels are often arbitrary or controversial. Instead they provide objective historical and social information on the languages, enabling users to decide for themselves which if any labels to apply. Accordingly, for present purposes, it seemed best to make use of the entire 76-language sample, bearing in mind that these may include languages with variegated historical and social profiles, a few of which are probably not appropriately considered to be creoles. The results of the survey are presented in Table 3.2:[6]

[6] For the most part, I assigned the feature values on the basis of the examples provided in the *APiCS* database. However, in a few cases, mostly those involving usual or obligatory core-argument flagging and

TABLE 3.2 Core-argument flagging and TAM marking in creoles (the *APiCS* languages)

		TAM marking		
		optional	obligatory	total
core-argument flagging	restricted or absent	68	6	74
	usual or obligatory	2	0	2
	total	70	6	76

The *APiCS*-language data in Table 3.2 bear a close resemblance to the sign-language data in Table 3.1, albeit with a small number of exceptions. Of the 76 languages in the *APiCS* sample, 74 have restricted or absent core-argument flagging, while 70 have optional TAM marking. Of the 76 languages, 68 share the same configuration of feature values as all of the 32 sign languages, with both restricted or absent core-argument flagging and optional TAM marking. Thus, the results of the survey provide additional empirical support for the suggestions, discussed earlier, that sign languages bear a typological resemblance to creole languages.

The data in Table 3.2 also provide further empirical support for claims to the effect that creoles differ in systematic ways from their non-creole counterparts. Whereas usual or obligatory core-argument flagging is widespread amongst non-creole languages, it is found in only 2 of the 76 *APiCS* languages. Similarly, while obligatory TAM marking is common in non-creole languages, occurring in 491 of the 868 languages in the Gil (to appear) sample, constituting 57%,[7] it is found on only 6 of the 76 *APiCS*

obligatory TAM marking (which could not be adequately inferred from inspection of a small number of examples), I made use of additional sources. The 76 *APiCS* languages, and the additional sources, where used, are as follows: African American English; Afrikaans; Ambon Malay; Angolar; Bahamian Creole; Batavia Creole; Belizean Creole; Berbice Dutch; Bislama; Cameroon Pidgin English; Cape Verdean Creole of Brava; Cape Verdean Creole of Santiago; Cape Verdean Creole of São Vicente; Casamancese Creole; Cavite Chabacano, Sippola (2011); Chinese Pidgin English; Chinese Pidgin Russian; Chinuk Wawa; Creolese; Diu Indo-Portuguese, Cardoso (2009); Early Sranan; Eskimo Pidgin; Fa d'Ambô; Fanakalo; Ghanaian Pidgin English; Guadeloupean Creole; Guinea-Bissau Kriyol; Gullah; Gurindji Kriol; Guyanais; Haitian Creole; Hawai'i Creole; Jamaican; Juba Arabic; Kikongo-Kituba; Kinubi; Korlai; Krio; Kriol; Lingala; Louisiana Creole; Martinican Creole; Mauritian Creole; Media Lengua; Michif; Mixed Ma'a/Mbugu; Negerhollands, van Rossem and van der Voort eds. (1996); Nengee; Nicaraguan Creole English; Nigerian Pidgin; Norf'k; Palenquero, Armin Schwegler (pc); Papiá Kristang; Papiamentu; Pichi; Pidgin Hawaiian; Pidgin Hindustani; Principense; Reunion Creole, Staudacher-Valliamée (2004); San Andres Creole English; Sango; Santome; Saramaccan; Seychelles Creole; Singapore Bazaar Malay; Singlish; Sranan; Sri Lanka Portuguese, Smith (1979); Sri Lankan Malay; Tayo; Ternate Chabacano, Sippola (2011); Tok Pisin; Trinidad English Creole; Vincentian Creole; Yimas-Arafundi Pidgin, William Foley (pc); Zamboanga Chabacano.

7 Note that of the 868 languages in the Gil (to appear) sample, 10 are also *APiCS* languages (Bislama, Cape Verdean Creole of Santiago, Haitian Creole, Jamaican Creole, Lingala, Papiamentu, Sranan, Sango, Seychelles Creole, and Sri Lankan Malay), while an additional language is a creole not included in the *APiCS* sample (Ndyuka). If these 11 languages are omitted from the 868-language sample, then out of the remaining 857 languages, 488 have obligatory TAM marking, which still comprises around 57%.

languages, or 8%.[8] In particular, whereas the combination of usual or obligatory core-argument flagging and obligatory TAM marking is common amongst non-creole languages, it is completely unattested amongst the 76 *APiCS* languages.

Nevertheless, as is the case for sign languages, creoles do not stand completely apart from other older languages with respect to the two features in question; rather, they occupy a proper subset of the typological space of non-creole languages. Consider the following example from Jamaican Creole:

(7) *Jamaican Creole* (Joseph T. Farquharson pc)
Di man si di uman
DEF man see DEF woman
'The man saw the woman'

With its absence of any core-argument flagging and TAM marking, Jamaican Creole sentence (7), like its Kata Kolok counterpart in (6), resembles similar sentences in many spoken languages. The difference between creole and non-creole languages is that few creole languages have either usual or obligatory core-argument flagging or obligatory TAM marking, two features that are characteristic of many non-creole languages.

In terms of complexity, the data in Table 3.2 suggest that creole languages do indeed tend to be simpler than their non-creole counterparts with respect to these particular features. However, the existence of many older languages sharing the same combination of feature values points towards a modified version of the creoles-are-simpler hypothesis. Whereas McWhorter (2001, 2005) argues for a bidirectional implication, 'simple if and only if creole' (or, equivalently, 'complex if and only if non-creole'), Gil (2001a, 2007) argues that this implication needs to be weakened to a unidirectional one, 'simple if creole' (or, equivalently, 'non-creole if complex'), the crucial cases being languages such as Riau Indonesian and others, which are as simple as creole languages without there being any independent external evidence for them ever having undergone any of the historical processes associated with creolization.[9] When set against the corresponding data from non-creole languages, the data in Table 3.2 provide further support for such a weakened uni-directional implication,

[8] Given the large amount of attention that TAM marking in creole languages has attracted, it may come as a surprise to find that for most creoles such marking is optional. Often, scholars are more interested in complex constructions to the detriment of their simpler counterparts. It is worth recalling Bickerton's (1977: 58–60) claim that creoles typically make use of unmarked verbs to express either past or present tense (depending on the semantic category of the verb), a construction which, in the present survey, is considered to be indicative of optional TAM marking.

[9] Although Parkvall (2008) also appears to be arguing for a bidirectional implication, his data seem to support the weaker unidirectional implication. Thus, out of 76 non-creole languages in his initial sample of 78, the simplest 5, Pirahã (0.18), Hmong Njua (0.20), Kobon (0.20), Maybrat (0.22), and Usan (0.23), are simpler than 13 out of his 31 creoles, namely Kinubi (0.24), Sranan (0.24), Guadeloupean (0.25), Nigerian Pidgin (0.25), Saramaccan (0.25), Seychellois (0.25), Guinea Bissau Creole (0.27), Jamaican (0.27), Palenquero (0.27), Krio (0.28), Haitian (0.29), Sãotomense (0.31), and Papiamentu (0.33). (The figures represent values on a scale from 0, maximally simple, to 1, maximally complex.)

showing that with respect to the two features under consideration, creole languages cluster towards the simpler end of the available spectrum, while not going beyond what is also possible for at least some non-creole languages.

In order to show that the relatively uniform profile of creole languages with respect to the flagging of core arguments and the marking of TAM categories is due to their being young languages, it is necessary to rule out alternative explanations appealing to characteristics of their source languages, which, due to historical accident, may constitute a skewed sample not representative of the world's languages. In principle, any property of a creole language may be attributed either to universal principles governing the ways in which creoles are formed, or to contingent features of the superstrate and/or substrate languages from which the creole in question is derived. Tables 3.3 and 3.4 below present an attempt to tease these possibilities apart.

In Tables 3.3 and 3.4, a selection of *APiCS* languages are compared with their source languages with regard to the flagging of core arguments and the marking of TAM. The first three columns present the *APiCS* language and its feature values, the next three columns its lexifier language and its feature values, and the final three columns the other language (or languages) that are generally assumed to have had a significant effect on the formation of the creole and its (or their) feature values.[10] For a 'classical' creole, the 'lexifier' and 'other' languages will correspond to superstrate and substrate respectively. However, in other cases, the notions of superstrate and substrate may be less appropriate, hence the more neutral terms 'lexifier' and 'other'.

Table 3.3 presents a selection of 10 out of the 68 *APiCS* languages with both restricted or absent core-argument flagging and optional TAM marking.

For some of the 68 *APiCS* languages, represented in Table 3.3 by Singapore Bazaar Malay and Ambonese Malay, their feature values are also shared by their lexifiers and other source languages. Hence the data provide no way to discriminate between universalist, superstratist and substratist explanations. However, for a larger number of the 68 languages, represented in Table 3.3 by the remaining 8 languages, the *APiCS* languages are simpler than their lexifier and/or source languages with respect to either one or both features. In such cases, then, a superstratist or a substratist account for the characteristics of the *APiCS* language can be ruled out accordingly. Thus, Haitian Creole and Batavia Creole are simpler than their lexifier languages with respect to TAM marking, while Chinese Pidgin Russian and Pidgin Hindustani are simpler than their lexifier languages with respect to both core-argument flagging and TAM marking. For these four languages, then, a superstratist explanation for their greater simplicity is ruled out (though a substratist explanation is not). Moving down,

[10] For the sake of brevity, the feature values are represented with the symbols 'Ø' and '+', where 'Ø' stands for the simpler feature value, namely, restricted or absent core-argument flagging and optional TAM marking respectively, while '+' stands for the more complex feature value, usual or obligatory core-argument flagging and obligatory TAM marking respectively.

TABLE 3.3 Creoles and their source languages: 10 selected *APiCS* languages with both restricted or absent flagging and optional TAM marking[11]

creole	flagging	TAM marking	lexifier	flagging	TAM marking	other	flagging	TAM marking
Singapore Bazaar Malay	Ø	Ø	Malay	Ø	Ø	Southern Min, Cantonese	Ø	Ø
Ambonese Malay	Ø	Ø	Malay	Ø	Ø	Maluku languages	Ø	Ø
Haitian Creole	Ø	Ø	French	Ø	+	Fongbe	Ø	Ø
Batavia Creole	Ø	Ø	Portuguese	Ø	+	Malay, Sundanese, Balinese	Ø	Ø
Chinese Pidgin Russian	Ø	Ø	Russian	+	+	Chinese	Ø	Ø
Pidgin Hindustani	Ø	Ø	Fiji Hindi	+	+	Fijian	Ø	Ø
Juba Arabic	Ø	Ø	Sudanese Arabic	Ø	+	Bari	Ø	+
Palenquero	Ø	Ø	Spanish	Ø	+	Kikongo	Ø	+
Diu Indo-Portuguese	Ø	Ø	Portuguese	Ø	+	Gujarati	+	+
Ternate Chabacano	Ø	Ø	Spanish	Ø	+	Tagalog	+	+

Juba Arabic and Palenquero are simpler than both their lexifier and their other source languages with respect to TAM marking, while Dui-Indonesian Portuguese and Ternate Chabacano are simpler than both their lexifier and their other source languages with respect to TAM marking and also simpler than their other source languages with respect to core-argument flagging. For these four languages, then, both superstratist and substratist accounts for the characteristics of the *APiCS* language can be ruled out. In summary, Table 3.3 suggests that with respect to the majority of *APiCS* languages

[11] Data for the non-*APiCS* languages in Table 3.3 are based on my own knowledge except for Fiji Hindi, (Moag 1977); Sudanese Arabic (Bergman 2002); Fijian (Schütz 1985); Bari (Nyombe 1987; Cohen 2000: 79); Kikongo (Laman 1912); and Gujarati (Mistry 1998; Doctor 2004).

with both restricted or absent core-argument flagging and optional TAM marking, superstratist and substratist explanations will not work. The only uniform account for their greater simplicity in comparison to other non-creole languages can be a universalist one, based on general principles governing the formation and development of creole languages.

Table 3.4 examines the 8 *APiCS* languages that are exceptional in that, unlike most other creole languages, they do not exhibit the usual pattern of restricted or absent core-argument flagging combined with optional TAM marking; with respect to these two features they are thus more complex than most other creoles.[12]

TABLE 3.4 Creoles and their source languages: All 8 *APiCS* languages with usual or obligatory flagging or obligatory TAM marking[13]

creole	flagging	TAM marking	lexifier	flagging	TAM marking	other	flagging	TAM marking
Lingala	Ø	+	Bobangi	Ø	+	neighboring African	Ø	+
Yimas-Arafundi Pidgin	Ø	+	Yimas	Ø	+	Arafundi	Ø	+
Sri Lankan Portuguese	Ø	+	Portuguese	Ø	+	Tamil	+	+
Reunion Creole	Ø	+	French	Ø	+	Malagasy, various Indian	+	Ø / +
Papiamentu	Ø	+	Portuguese, Spanish	Ø	+	various African	Ø	Ø / +
Sri Lankan Malay	Ø	+	Malay	Ø	Ø	Tamil	+	+
Zamboanga Chabacano	+	Ø	Spanish	Ø	+	Cebuano	+	+
Media Lengua	+	Ø	Spanish	Ø	+	Quechua	+	+

12 Interestingly, of these 8 languages, the one that happens to also occur in Parkvall's (2008) sample of 31 creoles, namely Papiamentu, is also the one that emerges as the most complex in his sample (see footnote 9 above). Similarly, in Gil's (2007:) experimental study of associational semantics, Papiamentu also turns out to be the most complex of the 3 creole languages included there. This suggests that at least some of the languages in Table 3.4 may share a somewhat more complex grammatical profile than most other creole languages, one that goes beyond the two grammatical features under consideration here.

13 Data for the non-APiCS languages in Table 3.4 are based on my own knowledge except for Bobangi (Whitehead 1899); Yimas (Foley 1991); Arafundi (William Foley pc); Tamil (Steever 2009,); Cebuano (Valkama 2000); and Quechua (Paul Heggarty pc).

These 8 languages do not appear to constitute a natural class with respect to any obvious properties pertaining to their histories and other external circumstances. Given the diverse nature of the original *APiCS* sample, one might expect to find a positive correlation between language age and complexity. Such a correlation would entail that the languages in Table 3.4 should be the oldest and most established of the creole languages, while on the other hand, pidgins would be the least likely to deviate from the usual pattern of restricted or absent core-argument flagging combined with optional TAM marking. However, the *APiCS* database does not support such a correlation (though it does not counterindicate it either). Of the 76 languages in the *APiCS* sample, 7 are characterized as having no (or virtually no) native speakers, a typical feature of young pidgins, and of these 7, one, namely Yimas-Arafundi Pidgin, actually appears in Table 3.4, due to its obligatory TAM marking.

Instead, what seems to be the case is that, for various reasons, each of the languages in Table 3.4 has inherited some aspects of complexity from one or more of its source languages. The first 6 languages exhibit obligatory TAM marking, which for Lingala, Yimas-Arafundi Pidgin, and Sri Lankan Portuguese is inherited from their lexifier and/or other source languages; for Reunion Creole and Papiamentu inherited from their lexifiers and/or some of their source languages; and for Sri Lankan Malay inherited specifically from its other source language. Similarly, the latter 2 languages exhibit usual or obligatory core-argument flagging, which for both Zamboanga Chabacano and Media Lengua is inherited from their other source languages. Because of such patterns of source-language influence, one may observe occasional instances of complexification where a creole emerges as more complex than either its lexifier or its other source languages. Thus, with respect to TAM marking, Reunion Creole and Papiamentu are more complex than those of their substrate languages which have optional TAM marking, while Sri Lankan Malay is more complex than its lexifier language. Similarly, with respect to core-argument flagging, both Zamboanga Chabacano and Media Lengua are more complex than their respective lexifier languages. Crucially, however, there are no instances of an *APiCS* language exhibiting greater complexity than *both* its lexifier *and* its other source language with respect to one or another of the two features under consideration: complexification is brought about only by inheritance from a source language. Not so simplification: even in Table 3.4, which deals with the most complex of the *APiCS* languages, instances of simplification are also in evidence. Thus, with regard to core-argument flagging, Sri Lankan Portuguese, Reunion Creole and Sri Lankan Malay are simpler than their substrate languages, while with respect to TAM marking, Zamboanga Chabacano and Media Lengua are simpler than both their lexifier and their other source languages. In balance then, even the 8 most complex *APiCS* languages show comparable amounts of complexification and simplification with respect to their various source languages. But while the instances of complexification can all be attributed to inheritance from their source languages, the instances of simplification reflect the general tendency for creoles to be simpler than other older languages.

In conjunction, then, Tables 3.3 and 3.4 show that while in certain cases *APiCS* languages may have inherited certain aspects of complexity from some of their source languages, the predominant pattern is one of simplification with respect to their source languages. Thus, the characteristic profile of creole languages as represented in Table 3.2, with their restricted or absent core-argument flagging and optional TAM marking, reflects their nature as young, newly-formed languages, which, in the relatively recent past, had to start from scratch.

3.5 Young languages and the development of predication

Sign languages and creoles thus share two important grammatical properties, namely, restricted or absent core-argument flagging and optional TAM marking. In this way, they display the simpler grammatical structures characteristic of newly-created languages that have not been around for long enough to accrue additional levels of complexity. However, these two properties are not unrelated. As suggested in section 3.2.3 above, they tend to correlate cross-linguistically, with restricted or absent core-argument flagging going with optional TAM marking, and usual or obligatory core-argument flagging being associated with obligatory TAM marking. In other words, these two properties of sign languages and creoles constitute two instantiations of a single deeper property. That deeper property, it is argued here, is the degree to which the grammar makes reference to predication.

3.5.1 *Predication*

In Gil (2012b) it is argued that in languages with usual or obligatory core-argument flagging and obligatory TAM marking, predication is strongly grammaticalized, playing a central role in morphosyntax and semantics. In contrast, in languages with restricted or absent core-argument flagging and optional TAM marking, predication may be more weakly grammaticalized with at most a minor impact on patterns of form and meaning. Thus, sign languages and creoles are languages in which predication tends to be weakly grammaticalized and may indeed be completely absent.

The way in which the notion of predication determines core-argument flagging and TAM marking is captured by the definition of predication provided in Gil (2012b) as a composite emergent entity derived from the alignment of two independent elements of conceptual structure: *thematic role assignment* and *headedness*. The definition takes as its point of departure an idealized language completely lacking in predication, the *Isolating-Monocategorial-Associational (IMA) Language* described in Gil (2005a).[14] In such a language, a basic two-part expression such as *CHICKEN*

[14] As defined in Gil (2005a), IMA Language bears a resemblance to Bickerton's (1990) *Protolanguage*, but differs from it in a number of respects, including the presence of recursion.

EAT would have a compositional semantics making reference to the Association Operator *A*, as in the formula: A (CHICKEN, EAT), which may be read 'entity associated with chicken and eat'. What this formula says is that the meaning of *CHICKEN EAT* must have something to do with chicken and eat; however, everything else is left underspecified. In particular, the expression is vague with regard to thematic roles, in the sense that the chicken could be either the agent or the patient of the eating. Indeed, in an appropriate context, it could bear some other thematic role. Similarly, the expression is indeterminate with respect to ontological category, in the sense that it could refer to an eating in which a chicken is involved (as in 'The chicken is eating') or to a chicken involved in an eating (as in 'the chicken that is eating'), or, once again, to meanings in other ontological categories such as time, place, manner or reason. In such a language, then, there is no predication; the semantic relationship between CHICKEN and EAT is entirely symmetric, as in the following representation:

(8) A (CHICKEN, EAT)

CHICKEN EAT

Representations such as in (8) may form the basis for the introduction of predication, via its two constituent components, thematic role assignment and headedness, as shown in the following semantic representations:

(9) (a) A (CHICKEN, EAT) *predication*

CHICKEN EAT

θ ⟵ ☜ *thematic role assignment*

head ☜ *headedness*

(b) A (CHICKEN, EAT) *attribution*

CHICKEN EAT

θ ⟵ ☜ *thematic role assignment*

head ☜ *headedness*

In both (9a) and (9b), EAT assigns a particular thematic role to CHICKEN. However, (9a) and (9b) differ with respect to headedness. In (9a) EAT is head and CHICKEN its modifier. What this means is that EAT projects its denotational identity up to the mother node, which accordingly denotes *eating associated with chicken*. Conversely, in (9b) CHICKEN is head and EAT its modifier. Here it is CHICKEN that projects its denotational identity upwards, resulting in *chicken associated with eating*. Thus, thematic role assignment and headedness reduce the vagueness of the original construction in (8). Specifically, the configuration in (9a) results in a predicative interpretation, commonly associated in English and other similar languages with clauses, while that in (9b) results in an attributive interpretation, typically associated in English and other such languages with NPs.

Predication is thus defined in terms of a particular alignment of thematic role assignment and headedness. A predicate is a head that assigns its modifier a thematic role. Conversely, an attribute is a modifier that assigns its head a thematic role.[15] This definition of predication differs from familiar ones in that it characterizes predication as a composite entity constructed from prior elements that are independently motivated features of conceptual structure, as is discussed in Gil (2012b). The imposition of thematic role alignment and headedness on a simpler structure lacking these elements is akin in spirit to the kind of 'Enrichment Schema' proposed by Jackendoff and Wittenberg (this volume, chapter 4) for other semantic domains.

The composite nature of predication suggests an explanation for the correlation between core-argument flagging and TAM marking. These two features may be viewed as readily observable proxies providing a measure of the degree to which thematic role assignment and tense are grammaticalized. While thematic role assignment is one of the constitutive elements of predication, tense too is arguably dependent on the other constitutive element, headedness. In Gil (2012b: 321–322) the relationship between tense and headedness is accounted for in terms of the dual verbal-cum-clausal character of tense, associated with the path of projection from the verbal head to the entire clause. Thus, by tying together thematic role assignment and tense, the composite notion of predication also brings together some of their formal expressions, such as core-argument flagging and TAM marking. Accordingly, languages with restricted or absent core-argument flagging and optional TAM marking are likely to be languages in which predication is at best only weakly grammaticalized and perhaps completely absent.

One such language, as argued in Gil (2001b, 2005a, 2012b), is Riau Indonesian. As illustrated earlier in (4), Riau Indonesian has restricted or absent core-argument flagging and optional TAM. Now consider the following naturalistic utterance:

[15] For ease of exposition, the above definition assumes binary branching. However, the more formal definitions of predication and attribution provided in Gil (2012b) allow for n-ary branching structures.

(10) *Riau Indonesian*
Gol cantik
goal beautiful
[watching replay of football goal on TV]
(i) 'That goal was beautiful'
(ii) 'That was a beautiful goal'

How to translate (10) into English? The grammar of English presents the translator with a dilemma, a forced choice between one of the two translations given in (10i) and (10ii). Whereas the former takes *cantik* 'beautiful' to be predicated of *gol* 'goal', in the latter *cantik* is assumed to be attributed of *gol*. These two translations thus point towards distinct potential grammatical analyses, involving distinct semantic and perhaps also syntactic structures. So which of the two translations is the right one—or is the sentence perhaps ambiguous between the two? The right answer is none of the above. Riau Indonesian provides little or no evidence for the grammaticalization of a predicate/attribute distinction. Moreover, in the actual utterance context, there is no reason to assume that the speaker had in mind one of these two translations to the exclusion of the other. Instead, the sentence is vague or unspecified with respect to the predicate/attribute distinction. It has a single unitary meaning which may be represented in terms of the association operator as A (GOAL, BEAUTIFUL), or, in plainer English, 'something to do with goal and beautiful'. In other words, the semantics of (10) is like that of the single representation in (8), not the two in (9). Examples such as (10), which are plentiful in any naturalistic corpus of Riau Indonesian, show how a language can enjoy an expressive power comparable to those of most familiar languages without recourse to predication.[16]

Work in progress suggests that Riau Indonesian is quite typical of colloquial varieties of Malay/Indonesian, and that other similar languages can be found in neighboring parts of Indonesia as well as other regions across the world. One suspects more such languages would be known if linguists were more sensitized to their possible existence. But obviously not all languages work this way. In particular, it is not the case that any word string in any language, bearing a superficial resemblance to (10), is amenable to the same kind of indeterminacy analysis. In many cases, such a string may instead represent an instance of ambiguity. Thus, the extent to which predication is grammaticalized in a given language can only be gauged though an in-depth examination of the morphosyntactic and semantic patterns of the language in question. With this qualification in mind, we now turn to examine the status of predication in sign languages and creoles.

[16] The present analysis of Riau Indonesian thus goes beyond that proposed by Jackendoff and Wittenberg (this volume, chapter 4) who, while acknowledging its relative simplicity in the syntactic domain, still presuppose predication in its semantics, a feature which, following most other scholars, they assume to be present in all languages.

3.5.2 *Predication in sign languages and creoles*

With their restricted or absent core-argument flagging and optional TAM marking, sign languages and creoles are predicted to have at best a weakly grammaticalized notion of predication or perhaps even none at all.

To test this prediction, a good point of departure is provided by the youngest of speech forms, those that have not yet acquired stabilized grammatical conventions or native speakers, and are perhaps not yet worthy of the label 'language'. One such speech variety is Israeli Geriatric Caretaker Jargon, used in a geriatric hospital ward by caretakers addressing patients suffering from dementia. While the administrative language and main lingua franca of the ward is Hebrew, a majority of the caretakers are recent immigrants whose native languages are primarily Russian and Amharic, while the patients are from a wide variety of different linguistic backgrounds, though their current linguistic abilities are generally significantly diminished. Thus, when speaking to the patients, the caretakers will choose words from either Hebrew or the patient's own native language, even though the caretaker may only be familiar with a small number of words or frozen expressions in the latter. The grammar will also be greatly simplified, because of the caretakers' frequent lack of knowledge of their patients' languages, and also because of their desire to accommodate their speech to the very limited remaining linguistic abilities of their patients. Since the situation is one of frequent turnover amongst both staff and patients, with different caretakers exhibiting very different speech patterns, there is little or no conventionalization specific to this particular speech variety, and hence it is most appropriately characterized as a jargon rather than, say, a pidgin.

The following is an utterance produced by a caretaker whose native language is Spanish speaking to a dementia patient whose native language is Russian:[17]

(11) *Israeli Geriatric Caretaker Jargon*

ugá	xaróša,	(kúšat	pažálsta)
Hebrew	Russian	(Russian	Russian)
cake:SGF	good[18]	(eat:INF	please)
cake	good	(eat	please)

[17] Example (11) is taken from a small opportunistic diary corpus compiled during visits to the Netanya Geriatric Medical Center. In (11), the first line presents a transcription in the presumed phonemic inventory of the speaker, the second line indicates the source language of each word, the third line provides an interlinear gloss of the word in the source language, while the fourth line presents a presumed interlinear gloss of the utterance within Israeli Geriatric Caretaker Jargon.

[18] Curiously, *xaróša* is a non-existent word in Russian. Its nearest available counterparts would be either the 'long' adjectival form *xaróšaja* 'good:NOM:SGF' or the corresponding less-commonly-occurring 'short' or 'predicative' form *xarašá* 'good:PRED:SGF'. One may conjecture that the speaker assumed a stem *xaróš-* and then invented a feminine suffix *-a*, given that in Hebrew, *ʕuga* 'cake' is feminine, and by happy coincidence *-a* happens to be a feminine marker in Spanish, Hebrew and Russian.

[native Spanish-speaking caretaker trying to get Russian-speaking patient to eat some cake]
(i) 'The cake is good, (please eat some)'
(ii) 'It's good cake, (please eat some)'

Focusing on the first part of the utterance, *ugá xaróša*, the question arises how to analyze it, whether as vague between predicative and attributive meanings, like Riau Indonesian (10), or as ambiguous. Since the utterance in (11) is not representative of a stable, conventionalized language, there are no grammatical rules that might provide the basis for a systematic distinction between predicative and attributive constructions. Moreover, the context is clearly one in which the speaker could plausibly be said not to care about this distinction. Thus, the most appropriate conclusion is that *ugá xaróša* involves neither predication nor attribution, but rather a simpler semantic structure such as that represented in (8), namely: A (CAKE, GOOD), or 'something to do with cake and good'.

Israeli Geriatric Caretaker Jargon may be typical of many other similar jargons and early pidgins, which might accordingly also be analyzed as lacking in predication. But what about sign languages and creoles, with their more stabilized grammatical systems that have already begun to accrue various kinds of linguistic complexity? The following are examples of simple two-word constructions in two sign languages:

(12) *Kata Kolok* (Connie de Vos pc)
DOG BIG
dog big
(i) 'The dog is big'
(ii) 'big dog'

(13) *Ugandan Sign Language* (Sam Lutalo-Kiingi pc)
WOMAN BEAUTIFUL
woman beautiful
(i) 'That woman is beautiful'
(ii) 'beautiful woman'

Both examples admit either predicative or attributive translations, as indicated above. Impressionistically, most or all sign languages seem to readily allow for constructions such as these. In the absence of sufficient information on the grammatical structures of these languages, and of naturalistic corpora with real live utterances embedded in communicative contexts, it is hard to say for certain whether such constructions are vague or ambiguous. However, in view of their widespread occurrence across sign languages, Occam's Razor would seem to suggest that, as a default analysis at least, such constructions should be assigned unitary semantic representations similar to that in (8), namely: A (DOG, BIG), 'something to do with dog and big' and: A (WOMAN, BEAUTIFUL) 'something to do with woman and beautiful'—that is to say, meanings that are vague with regard to the distinction between predication and attribution.

Spoken-language creoles, it would seem, present a somewhat less consistent picture. In some creoles, at least, it is hard or impossible to find constructions such as those in (10)–(13). Thus, some creoles, for example Afrikaans, exhibit subject-predicate and, in at least some cases, attribute-noun word order, thereby ensuring that predicates and attributes will appear on opposite sides of their sister constituents. Similarly, in some other creoles, such as Haitian Creole, the linear order of predicates and attributes relative to their sister constituents is the same, but predicative and attributive constructions may be marked differently, for example by the different positioning of an article. Nevertheless, some creoles do offer such constructions:

(14) *Juba Arabic* (Shuichiro Nakao pc)

timsâ	kebîr
crocodile:SG	big:SG

(i) 'The crocodile is big'
(ii) 'big crocodile'

(15) *Tok Pisin* (Hilário de Sousa pc)

Dok	i	ranim	pik
dog	PRED	run:TRNS	pig

(i) 'The dog is chasing the pig'
(ii) 'the dog that's chasing the pig'

Here too, as with the sign-language examples in (12) and (13), both predicative and attributive translations are available, and here too, the question arises whether such constructions are most appropriately analyzed as vague or ambiguous. In comparison to sign languages, however, the answer is probably less straightforward. In particular, in both cases above, one of the interpretations appears to be preferred over the other. In Juba Arabic (14), there is a preference for the attributive interpretation, since for the predicative one, the noun *timsâ* would be more likely to occur with an article. Conversely, in Tok Pisin (15), there is a preference for the predicative interpretation, since relative clauses tend to be used less commonly. Thus, the general picture is one in which creoles would appear to differ with respect to the degree to which predication is grammaticalized. While some spoken-creoles seem to provide more evidence for the grammaticalization of predication than any known sign language, many creoles still provide less evidence for its grammaticalization than a large number of older languages.

The problem when analyzing constructions such as those in (12)–(15) is that the necessary information for choosing between analyses is not readily available. We are often not in a position to be able to tell whether they are vague, like Riau Indonesian (10), or ambiguous. While some of the languages are underdescribed, with little information available overall, others are quite well documented, and yet it may still be quite difficult to extract the necessary information from the available descriptions. In part, at least, this is because most descriptive traditions and theoretical frameworks simply do not acknowledge the possibility that a construction might be unspecified for the distinction between predication and attribution, so deeply are these notions ingrained into our ways of thinking about language. Although at present precious few languages are described as permitting predication-attribution vagueness, I suspect that the number of such languages—old as well as new—might increase significantly once we become more open to the possibility of their existence.

In summary, then, there seems to be a tendency for younger languages to have less strongly grammaticalized predication than older languages. Moreover, within younger languages, there also appears to be a cline, with jargons and pidgins having the least amount of predication, sign languages perhaps a little bit more, and creoles probably significantly more, albeit still less than many other older languages. Though of course, within this overall picture, there may be exceptions, such as Riau Indonesian and other similar older languages with only weakly grammaticalized predication.

Whereas the traditional conception of predication as a primitive atomistic notion does not allow for any shades of grey, you either have it or you don't, the composite definition of predication proposed above allows for varying degrees of predication to be strung out on a continuum. Within a single construction, there may be differential preferences, whereby, for example, a predicative interpretation is preferred over a non-predicative one, but the latter is still available. Across constructions, there may be differences in the extent to which the predicative/attributive distinction is upheld. And from one language to another, the role of predication in the grammar may run the gamut from non-existent to very large. Yet another possibility is that predication may be unmarked in the grammar but still emerge as the product of frequency effects in discourse, as is argued in Gil (2012b) for Riau Indonesian, which is accordingly characterized as having ad hoc as opposed to systematic predication.

The emergent nature of predication makes it possible to talk about the connection between predication and language age. As suggested by the Israeli Geriatric Caretaker Jargon example in (11), languages start out without any systematic convergence of thematic role assignment and headedness, that is to say, without predication. As languages get older, discourse patterns associating thematic role assignment and headedness give rise to ad hoc predication, and then become conventionalized into morphosyntactic patterns, resulting in systematic predication. Sign languages and creoles occupy intermediate stages along this trajectory, which of course may differ idiosyncratically from language to language. Finally, it should be acknowledged that older languages may sometimes reverse course and simplify their inventory of grammatical rules, thereby moving back from systematic to ad hoc predication, as is the case for Riau Indonesian.

3.5.3 *Predication in ontogeny and phylogeny*

Sign languages and creoles represent just one way in which language can be young, but there are two other kinds of 'young language', namely, that of young infants and that of our hominin ancestors. Young languages such as sign languages and creoles may provide a model for early child language and also early hominin language, as has been suggested, for example, by Bickerton (1990) with respect to creoles and proto-language, and by Hurford (2012) with regard to sign languages and language evolution. Indeed, in the case at hand, the development of predication in sign languages and creoles may point towards a similar trajectory in both ontogeny and phylogeny.

Consider the following example from early English child language, also discussed by Jackendoff and Wittenberg (this volume, chapter 4):

(16) *English child language* (Brown 1973, quoting Lois Bloom)
Mommy sock
(i) 'Mommy's putting a sock on me'
(ii) 'Mommy's sock'

The above example is reported as representing two distinct utterances by the same child, in two different contexts. On one occasion it is translated as (16i), on the other as (16ii). The underlying assumption (which Jackendoff and Wittenberg share) seems to be that the string in question is ambiguous between these two different interpretations. Note that whereas the first interpretation is predicative, with *sock* as predicate (it might also have been translated as 'Mommy's socking me'), the second interpretation is attributive, with *Mommy* as attribute. Although failing to completely parallel (10)–(15), since it is not the same utterance that is either predicative or attributive, example (16) poses the by-now familiar analytical question: how do we know that it is ambiguous as opposed to merely vague? Without an extensive description of the child's grammatical competence at the age in which the two utterances were produced it is hard to provide a definitive answer. However, we should at least be open to the possibility that the child's grammar at the age in question assigned the same unspecified semantic representation to both utterances, namely: A (MOMMY, SOCK), or 'something to do with Mommy and sock', with only the different contexts of the respective utterances distinguishing between the two. More generally, it would seem reasonable to posit a stage in early child language where words are strung together without benefit of any kind of notion of predication—see Gil (2005a) for further discussion.

Ideally, this discussion would conclude with a similar example from an ancient hominin, but in the absence of time travel, such data remain out of our reach. Still, evidence from the linguistic behavior of great apes in captivity such as the bonobo Kanzi, documented by Greenfield and Savage-Rumbaugh (1990) and the orangutan Chantek, described by Miles (1990), suggests that apes have the ability to string words together and interpret such combinations compositionally, but lack the additional ability to impose further structural elements such as thematic role assignment and headedness, the prerequisites for predication—see Gil (2005a, 2006, 2012a). The ability to form simple compositional semantic representations such as that in (8) may accordingly be reconstructed to our common ancestors some 13 million years ago, and then assumed to have been actualized at some point down the line by one or more of our ancestor species.[19] To the extent that such simpler representations, lacking predication, are observable in sign languages, some creoles, and even some older languages today, they may be characterized as 'living fossils', in accordance with proposals put forward by Progovac (2006, 2009, this volume, chapter 5).

While ontogeny is often said to recapitulate phylogeny, one might add that typology, in particular that of sign languages and creoles, can also provide a useful window into both ontogeny and phylogeny. Thus, the development of predication in sign languages and creoles may provide a model for its development in early child language and the language of ancient hominins, with a similar course of grammatical and semantic complexification being followed by sign languages and creoles over decades, by young infants over months, and by our evolving ancestors over tens and perhaps hundreds of millennia. In turn, as is also suggested by Progovac (2006, 2009, this volume, chapter 5) and Jackendoff and Wittenberg (this volume, chapter 4), ontogeny and phylogeny will then find themselves reflected in the architecture of grammar, with later and more complex structures built on top of earlier and simpler ones.

Acknowledgments

A work like this could not come to fruition without the support of more of my colleagues than I have the space or the memory to thank. I am particularly grateful to Connie de Vos, Ray Jackendoff, John McWhorter, Nick Palfreyman, Ljiljana Progovac and Ulrike Zeshan for helpful discussions of issues addressed in this chapter; to Martin Haspelmath and Susanne Michaelis for allowing me access to a

[19] The approach outlined herein is thus akin to a recent suggestion within the minimalist framework by Hornstein and Pietroski (2009: 115) to the effect that 'I-language composition adds something to more ancient cognitive operations that are available to other animals', where 'I-language composition' corresponds to various more complex operations such as predication and 'more ancient cognitive operations' to the simpler semantic compositionality effected by the Association Operator. The present approach differs, however, from theirs and most other minimalist ones in acknowledging the possibility that contemporary human languages may vary significantly with respect to the degree to which they make use of the possibilities provided by 'I-language composition'.

pre-publication version of the *APiCS* database; to participants at the Seattle Workshop on Formal Linguistics and the Measurement of Grammatical Complexity and at various other seminars for many helpful comments; to the numerous language experts who patiently answered my detailed questions regarding their respective languages; and to Bernard Comrie, Fritz Newmeyer, Laurel Preston and Ulrike Zeshan for comments on an earlier version of this chapter.

4

What you can say without syntax: a hierarchy of grammatical complexity

RAY JACKENDOFF AND EVA WITTENBERG

4.1 Introduction

We present here the preliminary results of a large-scale thought experiment: how much and what kind of thought can be expressed without resorting to the tools provided by fully complex syntax? What can semantic structure alone accomplish before syntax is needed to regiment the message? How much expressive complexity can be created with minimal tools?[1]

These questions are more than just a mind game. We address them by developing a hierarchy of successively more complex grammars, beginning with trivial grammars that permit just one word per utterance, and culminating in fully complex grammars of natural languages, such as Dutch, Turkish, and Mohawk. The goal of the hierarchy is to provide a framework that accounts for differences in syntactic complexity across languages and language-like systems. But we find that it also can be used to describe certain subsystems of fully complex languages, stages of language acquisition, aspects of language processing, and, crucially, the interplay of language with other cognitive systems.

[1] For invaluable comments on many iterations of this material, we are deeply indebted to the members of the Tufts linguistics research seminar, past and present: Neil Cohn, Ari Goldberg, Anita Peti-Stantić, Rob Truswell, Stephanie Gottwald, Chelsey Ott, Naomi Berlove, Geert Booij, and Rabia Ergin. We also wish to thank Jesse Snedeker and her lab for important discussion, and David Gil, Dan Everett, Annie Senghas, Mark Aronoff, Susan Goldin-Meadow, Dea Hunsicker, Marie Coppola, Tom Bever, and Gina Kuperberg for answering our many questions about issues in their work that we are addressing here. Thanks also to Daniel Ross, Lila Gleitman, David Gil, Ljiljana Progovac, Fritz Newmeyer, and Laurel Preston for comments on a near-final version of this chapter. We are grateful for the opportunity to present this material not only at the University of Washington workshop on which the present volume is based, but also at the workshop on grammatical complexity at the University of Konstanz (2012), the University of North Carolina's 2012 Linguistics Colloquium, and the Tufts Conference on Language and Representation 2012.

Our model focuses on the mapping from meaning to sound and on the range of messages that different grammatical systems can convey. We assume that any differences in intelligence are negligible for people across the planet. They think the same thoughts, no matter what kind of grammatical system they use; and they express the same kinds of thoughts, regardless of the grammatical tools they have: past, present and future events, cause-and-effect relationships, social relationships, hypothetical questions, and so forth. Yet languages obviously differ in terms of their syntactic complexity.

Our idea is that the simpler grammars in our hierarchy put more responsibility for comprehension on pragmatics and discourse context. For instance, to understand a child's one-word utterance, one needs to rely heavily on inferences about what the child might have in mind. As the child's grammar acquires more grammatical devices, it provides more resources for making complex thoughts explicit, reducing the workload on the hearer. One could say that the syntactic complexity of a maturing speaker is gradually easing the semantic and pragmatic burden on the hearer. For another example of rich semantic and poor syntactic structure, consider Riau Indonesian, as described by David Gil (2005b, 2009, this volume, chapter 3). If a Riau speaker says *chicken eat*, the hearer has to figure out from context that the intended message is 'someone is eating that chicken over there' as opposed to, say, 'a chicken will eat that' or the many other things it could mean. Similarly, pidgins and creoles have syntactically simple ways of encoding complex thoughts that place high demands on the hearer.

The idea that grammars come with different degrees of complexity is not new. What is new, however, is our formal treatment of the semantic domain, weaving together context, word semantics, and combinatorial structure in a principled way. Unlike previous grammatical hierarchies, such as the familiar Chomsky hierarchy (Chomsky 1956), our goal is to create a model of increasingly complex sound-to-meaning mappings, not merely to generate sets of syntactically structured sentences. In our view, syntax is not the only generative linguistic system (Jackendoff 2002), so we challenge the often unspoken assumption that knowledge of syntax is the most essential component in knowledge of a language.

One old problem that falls by the wayside in our approach is the question of when children start 'having language.' In first language acquisition, they move on from very simple to more complex language, and there is no need to make a binary decision about their having language or not at any particular point in development. Similarly, no bright line is crossed when a pidgin (often treated as 'not a real language') develops into a creole (which *is* considered a language). Our proposed hierarchy serves as a formal tool for describing a range of communication systems, allowing us to ask what natural languages and what subsystems of natural languages fall where on the hierarchy. Thus the hierarchy offers a way to classify systems such as a child's language or a pidgin, not just as either language or not language, but rather as falling somewhere along a continuum of formal possibilities.

Our hierarchy can also serve a second, and theoretically stronger, purpose. To the degree that various points on the hierarchy are found to be actually instantiated, we will be able to make an empirical claim about the human language faculty: it is not a monolithic block of knowledge, but rather a palimpsest, consisting of layers of different degrees of complexity, in which various grammatical phenomena fall into different layers (see for example Progovac 2009 and this volume, chapter 5, for a parallel idea).

A few comments should be made about the limits and problems in such an endeavor. First, this is a work in progress, and we draw from limited data. Any generalizations beyond the data we have analyzed are tentative and waiting to be disproven. Second, we are asking readers to throw aside commonly held beliefs about the language system and follow us on a rough ride through very diverse areas of language. In particular, for the sake of argument, we are adopting the unfamiliar and sometimes painful methodology of assuming as little syntactic structure as possible. We hope to show that there is a lot to be gained, and not much lost, in adopting this methodology, and that in doing so, a wide variety of phenomena can be tied together.

Third, we should define what we mean by 'grammar', 'semantics', and 'syntax' here. In our terms, 'grammar' is the system that relates sound to meaning. It can include syntactic tools, but it need not: less complex grammars may be direct mappings of sound to meaning, without syntactic intervention. Semantics includes all the properties that have to do with meaning, linguistic and nonlinguistic, unless otherwise specified.[2] This definition has a large scope: among other things, it includes lexical semantics, information structural properties in a sentence such as Topic and Focus, and very general semantic concepts such as Object, Action, Agent, and Patient. Syntax, by contrast, is one step removed from meaning: it is a formal system that abstracts away from semantic categories, that labels constituents in terms of abstract categories such as nouns and verbs, and that imposes formal structure on them. Often, semantic and syntactic categories intersect. For instance, animate objects (a semantic category) are always nouns (a syntactic category). However, although actions (a semantic category) are usually verbs (a syntactic category), there are also action nouns such as *action* and *party*. A central point in our approach is that syntactic categories and rules are not automatically and inevitably needed in every linguistic system; instead of assuming their presence, our null hypothesis is that they are absent, and we check this hypothesis by performing standard tests for syntactic categories and phrases.

Since we are not assuming that syntax necessarily matches semantics, an important part of the grammar in our model is the system of *interface rules* that mediate between semantics, syntax, and phonology. As we will see, in these less complex grammars, certain aspects of the work traditionally assigned to syntax can be stated instead in terms

[2] In this chapter, we use the terms 'semantics' and 'pragmatics' rather informally: semantics, unless otherwise indicated, refers to meaning as a whole; pragmatics, to the aspects of meaning contributed by non-linguistic context and world knowledge.

of the way semantics maps into linguistic expression. (For more detailed exposition of our notion of grammar, see Jackendoff 2002; Culicover and Jackendoff 2005.)

Finally, we should emphasize that we are not proposing a global measure of linguistic complexity, such that one could decide, say, whether Warlpiri is more or less complex than ASL. We are inclined to think that there are many dimensions along which languages can be complex or not, and we are not sure that it is important to compare across dimensions. To take about the simplest possible case: what is more complex, a language with forty phonemes and two morphological cases, or a language with twenty phonemes and six cases? How would one decide? As Daniel Dennett puts it, speaking of biological complexity, "There is no single summit in Design Space, nor a single staircase or ladder with calibrated steps, so we cannot expect to find a scale for comparing amounts of design work across distant developing branches" (Dennett 1995: 134). Here we are exploring just one of those staircases, and in particular, steps rather near the bottom.

4.2 The hierarchy

This section introduces the hierarchy of grammatical complexity. On its own it is rather abstract. Its interest for us lies in what meanings can be expressed by the grammars that it defines, as will be seen in subsequent sections.

Logically, the simplest conceivable grammar is a **one-word grammar**, in which utterances are restricted to a single word, as in (1a). The next simplest is a **two-word grammar**, which allows utterances of either one or two words, as in (1b).[3] From there, it is a short way to a **concatenation grammar** (1c), which allows utterances to consist of word strings of arbitrary length. We state the grammar in the form of templates for (or constraints on) possible structures.

(1) a. **One-word grammar**
    $[_{\text{Utterance}}$ Word$]$ [*Traditional notation*: Utterance → Word]
  b. **Two-word grammar**
    $[_{\text{Utterance}}$ Word (Word)$]$ [Utterance → Word (Word)]
  c. **Concatenation grammar**
    $[_{\text{Utterance}}$ Word*$]$ [Utterance → Word*]

Notice that none of these grammars involves parts of speech or morphology. The relation of utterances to interpretations has to be conditioned by semantic distinctions, for example object vs. action, rather than by syntax, such as noun vs. verb. In

[3] Readers may wonder why we have singled out a two-word grammar as significant, but not, say, a three- or four-word grammar. From an empirical point of view, various phenomena we will be examining can be characterized in terms of two-word combinations. And from a theoretical point of view, moving from single words to pairs requires the speaker to manage not just the introduction of an additional word but also a semantic relation between them (see section 4.3.2). We speculate that this new semantic relation is the real source of complexity in two-word utterances.

other words, there is no real syntax in the usual sense, as defined by syntactic categories and syntactic markers of these categories.

The next elaboration, a **simple phrase grammar**, allows words to be grouped into phrases. An utterance consists of a concatenation of words and phrases, as in (2). But crucially, phrases cannot contain phrases, so embedding is limited to a depth of two nodes. Such a grammar requires two phrase structure rules, one for Utterances (2a) and one for Phrases (2b). There are in turn two variants of the rule for Phrases: one in which a Phrase consists only of two words (2bi), and one in which it may consist of an unlimited number of words (2bii).[4]

(2) **Simple phrase grammar**

a. [$_{\text{Utterance}}$ Word/Phrase*]	[Utterance → Word/Phrase*]
b. i. [$_{\text{Phrase}}$ Word Word] (2-word phrase)	[Phrase → Word Word]
ii. [$_{\text{Phrase}}$ Word*] (unlimited phrase)	[Phrase → Word*]

One way to introduce phrases is by prosody, resulting in a **prosodic simple phrase grammar**. In such a grammar, prosody indicates which words belong together semantically. For example, even without syntactic categories, a child's utterance like *kiss, Julius Daddy* might mean that one should kiss Julius's daddy, while *kiss Julius, Daddy* is more likely about Daddy kissing Julius himself.

Alternatively, at this point in the hierarchy it starts to become useful to introduce parts of speech (or syntactic categories) to label words and phrases, yielding a **part-of-speech simple phrase grammar**. In such a grammar, different categories of phrases may specify different categories and orders for the words they contain, and they may be marked differently: the word *kiss* would be translated into German as a verb with umlaut and a verbal suffix (*küssen*), but as a noun without the umlaut (*Kuss*).

In the next elaboration, *phrases* can be grouped into phrases, so that there is now the potential for recursion. (3) shows the general rules for a **recursive phrase grammar**: both Utterances (3a) and Phrases (3b) can consist of any combination of Words and Phrases.[5]

(3) **Recursive phrase grammar**

a. [$_{\text{Utterance}}$ Word/Phrase*]	[Utterance → Word/Phrase*]
b. [$_{\text{Phrase}}$ Word/Phrase*]	[Phrase → Word/Phrase*]

A different kind of elaboration of grammars involves the possibility of making words themselves composite, that is, compositional morphology. (4) shows the prototypes for two kinds of morphological rules.

[4] The notation *Word/Phrase** is meant to denote an unlimited concatenation whose elements are either Words or Phrases.

[5] Note that one-word, two-word, and concatenation grammars are among the simplest possible finite-state grammars, and simple and recursive phrase grammars are among the simplest possible context-free grammars. In our terms, the Minimalist Program's binary Merge (Chomsky 1995) amounts to a putatively universal recursive phrase grammar in which each nonterminal node has exactly two constituents.

(4) **Compounding:** $[_{\text{Word}}$ Word Word]
Affixal morphology: $[_{\text{Word}}$ {Word/Stem, Affix}] (either order)

For fully complex languages, we need to add further elaborations: functional categories (e.g. Det, Aux), principles governing long-distance dependencies, local binding of anaphors, perhaps grammatical functions, and so on.

As motivation for proposing these sorts of simple grammars, we note that a number of phenomena to be discussed here are described as having no subordination, no functional categories, little or no morphology, and semantically-driven principles of word order. In the proposed hierarchy, they all appear to be versions of concatenation grammar or simple phrase grammar.

4.3 Interface rules

4.3.1 *Interface rules for one-word grammars*

The classes of formal grammars proposed above, of course, are useful only when they are coupled with principles that state how overt utterances are linked to meanings—*interface rules.* There are two kinds of interface rules. The first kind is, simply, *words.* A word connects a piece of phonology to a piece of meaning. If it is used in a grammar that has syntactic features, a word also carries relevant syntactic features that allow it to connect with the syntactic structure. For example, the word *cat* is a linkage in long-term memory of a piece of phonological structure, some syntactic features, and some semantic features, as in (5).

(5) Phonology: $/\text{kæt}/_1$
Syntax: $[+\text{N}, -\text{V}, \text{count}]_1$
Semantics: $[\text{FELINE}, \text{PET}, \ldots]_1$

The subscripts indicate that these pieces of structure are connected and remain so when this unit forms part of a larger expression. Thus the word serves as part of the interface between phonology, syntax, and semantics.

The second kind of interface rule is a *combinatorial interface rule*, which specifies how the meanings of the parts of a grammatical constituent C are combined into the meaning of C. A simple case is adjectival modification of a noun: the adjective denotes a property of the entity denoted by the head noun. We will be dealing here mainly with this kind of interface rule.

An important generalization we have discovered is that the interface rules for lower levels of the hierarchy scale up to the upper levels. That is, they apply whether the constituent C is an Utterance, a Phrase, or a Word, whether its parts are Phrases, Words, or sub-word morphemes, and whether Phrases and Utterances are defined only prosodically or in terms of syntactic categories. This means that as one moves up the hierarchy, principles from lower points are not obliterated; rather they are elaborated. We will illustrate this point as we go along.

A one-word grammar (1a) presents the very simplest interface rule, (6): the meaning of the word equals the meaning of the utterance. In this rule and those to follow, the subscripts indicate links between phonology, syntax, and semantics. Here, the phonological Word, with subscript 1, is linked to the meaning with subscript 1, but this meaning is also linked to the entire Utterance.

(6) **Word = Utterance**
Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1]_2$
Semantics: $X_{1,2}$

Since utterances in this grammar contain only a single word, there is in fact no real distinction between words and utterances. So an even simpler form of this rule is the trivial (7).

(7) Phonology: Utterance_1
Semantics: X_1

(7) is the only interface rule necessary for most primate calls, which are essentially monadic utterances. What makes (7) more interesting is that it also accounts for the use of certain words of fully complex languages, words like *hello, ouch, upsey-daisy*, and *abracadabra*. Such words have no syntactic properties. Rather, they serve as full utterances on their own, and they appear in combination only in paratactic situations like *Hello, Bill* and quotative contexts like *"Hello," she said*, in which anything can be inserted, even a phrase from another language.

However, a one-word grammar need not be quite this limited. We can add a little more expressivity by allowing the utterance to paste some content around the meaning of the word, as shown in (8).

(8) **One-word grammar with pragmatics**
Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1]_2$
Semantics: $[\text{F}\ (\text{X}_1)]_2$

Here, the function *F* in the semantics is present in the meaning, but it has no overt counterpart in the utterance. (8) appears to characterize the one-word stage of child language. For instance, *doggie* can be used to mean 'there's the doggie,' 'where's the doggie,' 'that looks like a doggie,' 'that belongs to the doggie,' 'I want the doggie,' 'doggie, pay attention to me,' and so on, where each of these extra pieces of meaning is a pragmatically determined function *F*.

Pragmatic enrichment as formalized in (8) is not confined to one-word grammars. As promised above, it scales up to syntactically more complex languages, which make use of a more general version of rule (8) that we call the **enrichment schema**. It has two cases, one for one-phrase utterances (9), and one for coercions (11). (9) represents the general form of principles for interpreting one-phrase utterances, for example those in (10).

(9) **Enrichment schema:** One-phrase utterances
Syntax: $[_{\text{Utterance}}\ \text{Phrase}_1]_2$
Semantics: $[\text{F}\ (\text{X}_1)]_2$

(10) An eagle!!
Some scotch?
[What kind of pizza do you want?] – Pepperoni.

(11) represents the general form of coercions, such as those in (12). The underlined parts of the paraphrases in (12) spell out the unexpressed function *F*.

(11) **Enrichment schema:** Coercion
Syntax: $[_{\text{Phrase}}\ \text{Phrase}_1\ ]_2$ *or* $\text{Phrase}_{1,2}$
Semantics: $[\text{F}\ (\text{X}_1)]_2$

(12) Plato [= '<u>book by</u> Plato'] is on the top shelf, next to Chomsky.
The ham sandwich [= '<u>the person with</u> ham sandwich'] wants more coffee.
I'm [= 'my <u>car</u>'s'] parked out back.
(Sluicing) Joe ate something, but I don't know what [= 'what <u>he ate</u>'].

Notice that (8), (9) and (11) are identical except for the labels on the syntactic constituents. The basic principle for each case is that a constituent has extra added meaning beyond the meanings of its words. The literal meaning of the constituent, determined just by the meanings of the words, is semantically embedded in a larger, pragmatically determined meaning. Hence (8), the interface principle for one-word grammars, scales up to situations in more complex languages.

4.3.2 *Basic interface rules for 2-word grammars*

Next, let us amplify the grammar to two constituents that together form an utterance. Now the problem arises of how to combine their meanings to form the meaning of the whole. One thing that makes two-word utterances harder than one-word utterances is that it is necessary to establish an unspoken semantic relation between them. The simplest way to do this is just to say that the two meanings combine somehow or another. But this results in what Anderson (2004) calls 'semantic soup,' in which there is far too much indeterminacy for the hearer.[6] More discipline is necessary. To

[6] Gil's Association Operator (this volume, chapter 3) has this effect, in that the meaning of a constituent [X Y] is simply 'entity associated with X and Y.' A real example of 'semantic soup' can be found in the utterances of sign-trained apes, for example Nim Chimpsky's *give orange me give eat orange me eat orange give me eat orange give me you* (Seidenberg and Petitto 1978; Terrace 1979). Gil's notion of Predication

ground our solution intuitively, we illustrate not with actual two-word grammars, but with a subsystem of English that approximates a two-word grammar: compounds. In English compounds, the semantics can map directly into phonology: a sequence of two words.[7] (We discuss children's two-word grammars in a moment.)

There are at least three ways the constituents of a compound can combine semantically. First, one component can be a semantic argument of the other, as in (13). Examples are shown in (14).

(13) **Function-argument schema**
Phonology: $[_{\text{Word}}\ \text{Word}_1\ \text{Word}_2\ ]_3$
Semantics: $[F_2\ (X_1)]_3$
(exchange subscripts 1 & 2 for left-headed compounding)

(14) union member [= 'member of a union']
helicopter attack [= 'attack by helicopter']

Notice that the interface rule determines the word order, in that the semantic head, subscripted 2, corresponds to the second word (in English). Thus, syntactic structure is not needed to specify headedness; word order alone is enough. This illustrates a more general point of our approach mentioned in section 4.1: not all 'grammatical' effects arise from the syntax. Many are a consequence of the interface rules.

A second way that two-word meanings can combine is for one of the constituents to be a modifier of the other, as in (15), where the semicolon symbolizes the semantic relation of Y modifying X. Examples are shown in (16).

(15) **Modification schema**
Phonology: $[_{\text{Word}}\ \text{Word}_1\ \text{Word}_2]_3$
Semantics: $[X_2;\ Y_1]_3$

(16) blackbird [= 'bird that is black']
plastic container [= 'container that is plastic']

Often the interpretation of a compound involves enrichment in the sense of rule (11), in addition to modification. For instance, the constituents of the compounds in (17) are enriched by the underlined material before being combined by the modification schema.

corresponds to our Function-argument schema ((13) below); his Attribution corresponds to our Modification schema (15).

[7] Alternatively, the semantics might map into morphosyntax. However, English compounding is not particularly restrictive about the syntactic categories of the compound's constituents, as can be seen from examples like (i); and though compound nouns are predominant, compound verbs and adjectives are also possible, as seen in (ii).

(i) $\text{long}_A\ \text{bow}_N$, $\text{under}_P\text{current}_N$, $\text{pull}_V\text{over}_P$, $\text{over}_P\text{kill}_N$, $\text{speak}_V\text{easy}_A$, $\text{hear}_V\text{say}_V$, $\text{once}_?\text{over}_P$
(ii) $[\text{skin}_N\ \text{deep}_A]_A$, $[\text{worm}_N\ \text{eaten}_V]_A$, $[\text{blue}_A\ \text{green}_A]_A$, $[\text{house}_N\ \text{sit}_V]_V$

(17) snowman [= 'simulation of man that is made of snow']
garbage man [= 'man who takes away garbage']

A third way for two words to combine semantically is for them both to be arguments of an unexpressed function. (18) is the rule, and (19) gives an example.

(18) **Co-argument schema**
Phonology: $[_{\text{Word}}\ \text{Word}_1\ \text{Word}_2\ ]_3$
Semantics: $[\text{F}\ (\text{X}_1, \text{Y}_2)]_3$

(19) seahorse [= 'something that looks like a horse and that lives in the sea']

Notice that the form of phonology in these three schemas is the same; what differs is how it maps into semantics.

In two-word child language, we find the same interface principles at work, except that instead of mapping the semantics to a compound word consisting of a pair of words, they map the semantics to a complete utterance consisting of a pair of words, as in (20). (21) gives an example of each of the principles.

(20) Phonology: $[_{\text{Utterance}}\ \text{Word}_1\ \text{Word}_2]_3$

(21) Function-argument schema: *Mommy fix* [= 'Mommy should fix it']
Modification schema + enrichment: *Big house* [= 'that's a big house']
Co-argument schema: *Mommy pumpkin* [= 'Mommy cuts the pumpkin']

Roger Brown (1973), quoting Lois Bloom (1970), cites two instances of the same child uttering *mommy sock* (see also the discussion in Gil, this volume, chapter 3). On one occasion it apparently is intended to mean 'Mommy's sock,' a case of modification. On another occasion it is intended to mean 'Mommy's putting a sock on me,' a case of co-arguments.

Martin Braine (1963) observed that many (but not all) children enter the two-word stage with 'pivot schemas': a small number of words that occur in combination with a larger variety of other words, as illustrated in (22). This observation has been revived in recent years (Tomasello 2003).

(22) see baby, see pretty, etc.
more car, more cereal, more cookie, more fish, etc.
no bed, no down, no fix, etc.

In our approach, each pivot can be regarded as an interface rule, for example (23), which is a special case of the function-argument schema.[8]

[8] Thus we need not speak of 'pivot *grammar*,' as Braine did, implying that the child's language is made up entirely of pivot schemas. Rather, pivot schemas are just one possible component of the child's language.

(23) Phonology: $[_{\text{Utterance}}\ [_{\text{Word}}\ \text{see}]\ \langle \text{Word}_i \rangle\ ]_j$
Semantics: $[\text{SEE (ME/YOU, } (Y_i)\ ]_j$

Like the enrichment schema, the interface rules in this section apply also to phenomena in more complex languages. Compounding is not the only instantiation. The function-argument schema scales up to the familiar principles for integrating a syntactic head with its complements, and the modification schema scales up to the usual rules for integrating syntactic adjuncts with their heads. The co-argument schema scales up to, among other things, the English causal paratactic conditional (24), where there is no overt connective, and small clause constructions (25), which have an implicit copula that links the NP and the predicate (cf. Progovac, this volume, chapter 5).

(24) You shoot a cop, you go to jail.

(25) a. Everyone out of the car!
b. John at a baseball game?! (I can't believe it!)
c. [John at a baseball game] is hard to imagine.
d. No dogs allowed.
e. Refreshments in the kitchen.

Pivot schemas too have counterparts in complex languages. (26a) appends a free choice of name to a title to form an NP; (26b,c) combine a pivot with a free choice of phrase to form a full root utterance that cannot embed.

(26) a. *Forms of address*: Mr. X, Ms. Y, Governor Z, Rabbi W, etc.
b. *Directed epithets*: Fuck/Damn/The hell with/Hooray for NP!
c. *Inquiries*: How about XP? What about XP?

4.3.3 *Correlating word order with thematic roles*

The interface rules we have just described still leave the interpretations of many configurations indeterminate, in danger of ending up with uninterpretable 'semantic soup.' For instance, if we encounter the utterance *chicken eat*, we may infer that it is to be interpreted by the function-argument schema, with *chicken* as the argument of *eat*. But *which* argument? Is the chicken eating or being eaten? Similarly, the utterance *cow horse* might be interpreted by the co-argument schema to denote an action involving a cow and a horse. But which one is acting on which? The indeterminacy becomes even greater when we move to a concatenation grammar, which allows more than two words. For instance, does *cow big horse* mean that the cow is big, or the horse?

One widespread strategy to ameliorate this problem is well-documented both in the languages of the world and in the various less complex phenomena we have been looking at: the preference for Agents to precede Patients, so that the cow is interpreted as doing something to the horse, not vice versa (see Gil 2005b, for a similar discussion). This preference can be stated as a simple interface rule (27), which relates thematic roles to linear order, without any reference to further syntactic structure. In a two-word grammar, the syntactic part of the rule is (27a). By allowing the two words in question to be non-adjacent, the rule generalizes to a concatenation grammar, as in (27b). Notice that the semantic function *F* that assigns the thematic roles has no index in this particular rule, so the rule does not depend on the exact relation between Agent and Patient to be expressed. (The ellipses in the semantics allow for further unexpressed semantic information as well.)

(27) **Agent** > **Patient** (a special case of the Co-argument schema)

a. *Version for two-word grammar*

Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1\ \text{Word}_2]_3$

Semantics: $[\text{F (Agent: X}_1\text{, Patient: Y}_2, \ldots)]_3$

b. *Version for concatenation grammar*

Phonology/syntax: $[_{\text{Utterance}} \ldots \text{Word}_1 \ldots \text{Word}_2 \ldots]_3$

Semantics: $[\text{F (Agent: X}_1\text{, Patient: Y}_2, \ldots)]_3$

Another widespread strategy is for Agents to precede Actions. The formalization of this strategy for the case of concatenation grammars is stated in (28). It is of course a prototype for subject-verb order in more complex languages.

(28) **Agent** > **Action** (special case of the Function-argument schema)

Phonology/syntax: $[_{\text{Utterance}} \ldots \text{Word}_1 \ldots \text{Word}_2 \ldots]_3$

Semantics: $[\text{F}_2\ (\text{Agent: X}_1, \ldots)]_3$

Actions may precede Patients, as in (29), or vice versa, as in (30). (29) is the prototype for VO order; (30) for OV order. The choice between them is arbitrary, but whichever way the order is conventionalized in a particular system, a relatively stable order makes communication of meaning more reliable and less dependent on context.

(29) **Action** > **Patient** (Prototype for VO order)

Phonology/syntax: $[_{\text{Utterance}} \ldots \text{Word}_1 \ldots \text{Word}_2 \ldots]_3$

Semantics: $[\text{F}_1\ (\ldots, \text{Patient: X}_2, \ldots)]_3$

(30) **Patient** > **Action** (Prototype for OV order)

Phonology/syntax: $[_{\text{Utterance}} \ldots \text{Word}_1 \ldots \text{Word}_2 \ldots]_3$

Semantics: $[\text{F}_2\ (\ldots, \text{Patient: X}_1, \ldots)]_3$

Another strategy correlates linear order with information structure. Most commonly, topic comes first and focus last, as in (31) and (32).

(31) **Topic first**

Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1 \ldots]_2$

Information structure: Topic_1

(32) **Focus last**
Phonology/syntax: $[_{\text{Utterance}} \ldots \text{Word}_1\ ]_2$
Information structure: Focus_1

The possible strategies in (27)–(32) lead to the possible interpretations of *chicken eat* and *eat chicken* shown in (33), depending on which principles the system in question makes use of.

(33) a. *Chicken eat* = 'chicken is eating' by Agent > Action
Chicken eat = 'someone is eating chicken' by Patient > Action (if the language has it)
Chicken eat = 'someone is eating chicken' by Topic First (if the language has it)
b. *Eat chicken* = 'someone is eating chicken' by Action > Patient (if the language has it)
Eat chicken = 'chicken is eating' by Focus Last (if the language has it)

Like the function-argument, modification, and co-argument schemas of the previous section, the interface principles in (27)–(32) map directly between a string of phonological words and a meaning. They invoke only linear order and semantic distinctions such as object vs. action, argument vs. modifier, agent vs. patient, and topic vs. focus. They show how a fairly expressive language could be constituted without syntactic categories and even without phrase structure.[9]

At this point, one might argue that the correlation of word order with thematic roles is nothing but a 'perceptual strategy' or a 'habit'; it is not *real* language, which requires true syntax (Fodor, Bever, and Garrett 1974; Townsend and Bever 2001). We find this distinction questionable. Visually perceiving an event's participants in a certain order or with a certain prominence might be a perceptual strategy (Hafri, Papafragou, and Trueswell, 2013). But encoding the thematic roles in an utterance in terms of word order is a distinctly linguistic principle: it is part of the mapping between meaning and sound. In other words, even if principles like (27)–(32) might be grounded in more general cognition, they are thereby no less linguistic.

As mentioned above, Agent > Action, Agent > Patient, Action > Patient, and Patient > Action are models for SVO and SOV order in more complex languages. Moreover, Topic is often marked by initial position, while Focus is often marked by final or near-final position. So again the rudimentary interface principles for

[9] Jackendoff (in preparation) discusses two further important interface principles. The first, which applies in concatenation grammars and upward, is that a semantic constituent preferably corresponds to a contiguous string in phonology. This might be considered a precursor of syntactic phrase structure. The second, which appears in prosodic simple phrase grammars and upward, is that a prosodic phrase generally corresponds to a semantic constituent; a suitably nuanced version of this principle appears in more complex languages as well, though it is usually stated in terms of a correspondence between phonological and *syntactic* constituency (e.g Selkirk 1984, 2000; Jackendoff 1987, 2007).

concatenation grammars scale up to widespread principles in syntactically complex languages. The simpler principles don't disappear as we move up the hierarchy.

4.4 Illustrations of the hierarchy

4.4.1 *Some phenomena that involve lower levels of the hierarchy*

We now briefly mention some phenomena that appear to instantiate lower levels of the hierarchy (see Jackendoff, in preparation, for more detailed discussion).

Pidgins and creoles. Pidgins are often described (e.g. Bickerton 1981, 2008; Givón 2009; see also Gil, this volume, chapter 3) as having no subordination, no morphology, no functional categories, free omission of arguments, and unstable word order governed by semantic/pragmatic principles such as Agent First and Focus Last. Our outlook leads us to ask: is there any evidence in pidgins for parts of speech? Is there any evidence for phrasal categories such as NP? Is there any phrase structure at all aside from prosody? If not, pidgins would be classified as concatenation grammars—or if prosody is doing some work, they could be prosodic simple phrase grammars. Creoles, of course, do add many features of more complex languages, such as more conventionalized word order, functional categories, and syntactic subordination. In our terms, the transition from pidgin to creole is not from non-language to language, but rather one or more steps up the grammatical hierarchy.[10]

Late second language acquisition. Klein and Perdue (1997) describe a distinct stage in late second language acquisition that they call the Basic Variety. In their multi-language longitudinal study, they found that all learners achieved this stage; many speakers went beyond this stage, but many did not. The Basic Variety is described as lacking inflectional morphology and sentential subordination, and freely omitting arguments. It has simple, largely semantically based principles of word order including Agent First and Focus Last. (The relative position of Patient and Action seems to be determined by the target language.) Our framework leads us to ask the same questions for Basic Variety as for pidgins: is there any evidence for parts of speech or phrasal categories? And again, is there any phrase structure aside from prosody? We conjecture that the Basic Variety too is either a concatenation grammar or, if prosody is doing any work, a simple phrase grammar.

Home sign. A third case of a language with less complex syntax is home sign, the languages invented by deaf children who have no exposure to a signed language. These have been studied extensively by Susan Goldin-Meadow and colleagues for

[10] We recognize, of course, that there is considerable variation among pidgins and creoles. A description in our terms requires considerable care, differentiating one case from the next.

decades (e.g. Feldman, Goldin-Meadow, and Gleitman 1978; Goldin-Meadow 2003). On their account, they have at most rudimentary morphology and freely omit arguments. There are morphological differences between nouns and verbs (but in our terms, the distinction may in fact be between objects and actions). There appears to be no use of prosody to delineate semantic constituency. Homesigners do produce some sentences with multiple verbs (or action words), which Goldin-Meadow describes in terms of syntactic embedding. We conjecture that these are actually rudimentary serial verb or serial action-word constructions, which need not involve embedding.[11] Our current diagnosis of home signs is that they have only a semantic distinction of object vs. action, not a syntactic distinction of noun vs. verb, and that they have a concatenation grammar with a possible admixture of a small amount of semantically-based morphology.

Al-Sayyid Bedouin Sign Language. This is a language emerging in a Bedouin tribe in Israel with three generations of hereditary deafness (Sandler, Meir, Padden, and Aronoff 2005; Aronoff, Meir, Padden, and Sandler, 2008). Based on the very small amount of published data, the language of first generation signers looks like a one-word grammar with a slight admixture of two-word grammar. This form of the language places an extremely heavy reliance on context and pragmatics for understanding. The language of older second generation signers looks like a concatenation grammar, with little consistent use of prosody to delimit semantic constituents. The language of younger second generation signers looks like a simple phrase grammar in which prosodic constituency plays a grammatical role. There is still no morphology, and no evidence for parts of speech. In other words, as the language has developed, it has gradually climbed up the hierarchy. From what we can tell from the literature on Nicaraguan Sign Language (Kegl, Senghas, and Coppola 1999; Senghas, 2003), it has undergone a similar development, though evidently climbing up the hierarchy a good deal faster.

Processing strategies. The use of rules like (27–32) is not confined to emerging languages. Townsend and Bever (2001) discuss what they call semantically based 'interpretive strategies' or 'habits' that influence language comprehension and lead to garden path situations. In particular, hearers tend to rely on semantically based principles of word order such as Agent > Patient, which is why they have more difficulty with constructions such as reversible passives and object relatives. Similarly, Ferreira and Patson (2007) discuss what they call 'shallow parsing' or 'good-enough parsing' in sentences like *Bill knows the answer is wrong*: subjects in their experiments apparently rely on linear order and semantic plausibility rather than syntactic

[11] Hunsicker and Goldin-Meadow (2012) propose that one homesigner they have analyzed uses a syntactic NP constituent consisting of a demonstrative that functions as a determiner plus a content word that functions as head noun. Jackendoff (in preparation) shows how their evidence can be accounted for in terms of a concatenation grammar.

structure. As is well known, similar symptoms appear in language comprehension by Broca's aphasics (Caramazza and Zurif 1976). Dick *et al.* (2001) show that college students behave like Broca's aphasics when trying to understand sentences that have been low-pass filtered and sped up. As mentioned above, even though the literature tends to describe these so-called 'strategies' or 'heuristics' as something separate from language, they are still mappings from phonology to meaning—just simpler ones. Building in part on evidence and analysis in Kuperberg (2007), we propose that the language processor makes use both of rules of full syntactic grammar and of rules of simpler concatenation grammar. When the two types of rules produce conflicting analyses, interpretation is more difficult. And when syntactic rules break down under conditions of stress or disability, the concatenation grammar is still at work.

4.4.2 *Two syntactically less complex languages*

We now offer a brief synopsis of our work (Jackendoff, in preparation) on two languages whose full grammar appears not to make use of the entire hierarchy. One is **Riau Indonesian,** a vernacular with several million speakers, described in detail by David Gil (Gil 2005b, 2009, this volume, chapter 3, and many other publications). Gil argues that this language presents no evidence for syntactic parts of speech. There is a small number (less than 20) of affixes which are completely unselective as to what they attach to. There is no inflectional morphology. Arguments can be freely omitted. There is a small number (less than 30) of closed-class items which require complements; however, they are completely unselective except on semantic grounds (some select actions, some objects). The only evidence for constituent structure comes from prosodic phrasing. The effects expressed by syntactic subordination in English are expressed in this language through syntactic parataxis plus pragmatic enrichment (e.g. the enrichment schema in (9) and (11) and a version of the co-argument schema in (18)). For instance, conditionals are expressed only by a counterpart of the English paratactic conditional, e.g. *You shoot a cop, you go to jail* (24).

The word order in this language is substantially free, but heads tend to precede modifiers, agents tend to precede actions, and actions tend to precede patients. In addition, there are information structure influences, in particular a tendency for topics to appear at the beginning of a sentence. The only rigid word order is imposed by the closed class items, which require their complements on their immediate right. The syntactic and pragmatic freedom of this language can be illustrated by Gil's example (34).

(34) *ayam makan*, 'chicken eat' can mean
   a. '{a/the} chicken(s) {is/are eating/ate/will eat} {something/it}'
   b. '{something/I/you/he/she/they} {is/are eating/ate/will eat} {a/the} chicken'
   c. '{a/the} chicken that {is/was} eating'
   d. '{a/the} chicken that {is/was} being eaten'
   e. 'someone is eating with/for the chicken'
   f. 'where/when the chicken is eating'

Our analysis is that the language is basically a simple phrase grammar whose constituency is determined by prosody, with a small amount of morphology. The details of our analysis are presented in Jackendoff (in preparation) and are beyond the scope of this chapter. But basically, Riau Indonesian is a language that is syntactically simple in our sense.

The second example of a syntactically simple language is the controversial case of **Pirahã**, studied extensively by Dan Everett (Everett 2005, 2009). This language has exuberant suffixal morphology as well as compounding, so it is far from simple in this respect. Everett's analyses, in all their various incarnations, assume a noun-verb distinction, so the parts of speech may be at least noun, verb, and everything else. But there are no definite or indefinite articles, no markers of plurality, and no inflectional morphology. Like many of the cases we have looked at, arguments can be freely omitted, though there is also the possibility of pronominal clitics on the verb. There is fairly fixed SOV word order, with the possibility of postposing the subject and/or the object. There is an interesting restriction, namely that nouns appear to admit only one modifier, either a prenominal possessive or a postnominal adjectival or quantifier, but not both. We take this as tentative evidence for an NP constituent that is constrained to at most two words.

Everett's most notorious claim is that Pirahã lacks recursion. His evidence is that all the constructions expressed by recursive syntax in English are either paratactic or require circumlocution. In a critique of Everett (2005), Nevins, Pesetsky, and Rodrigues (2009) show that there may be multiple clauses in a sentence. However, they do *not* show that clauses can contain clauses that contain clauses, that is, true recursion. And in fact the multiple clauses they cite are arguably paratactic; so what can be shown at most is that Pirahã has a simple phrase grammar, with the only further depth of embedding being the two-word NP, essentially the same point in the hierarchy as Everett claims. In his response to the critique (Everett 2009), Everett makes the more radical claim that these clauses are separate sentences, so the language can be described as a concatenation grammar, possibly with NPs. So whichever analysis we adopt, Pirahã looks like a syntactically simple language, though not as simple as Riau Indonesian, in that it has syntactic categories and rich morphology.

4.5 Implications for the evolution of language

We are reluctant to get involved in debates on the evolution of the human language capacity, because there is no evidence for who was saying what when among our hominid ancestors, not to mention among our cousins the Neanderthals. However, the grammatical hierarchy developed here offers the possibility of speculatively reverse-engineering evolution: it offers possible language-like systems that are more complex than primate calls but less complex than modern full languages. We find it plausible that such systems existed in earlier hominids, and that further evolutionary steps resulted in brain structures that permitted the more complex languages that humans speak today. (Of course, finding something plausible is the weakest of arguments, but unfortunately that is for the most part the best one can do in debates on evolution of language.)

This is essentially the idea behind Derek Bickerton's (1990, 2008) 'protolanguage' hypothesis, further elaborated by Jackendoff (1999, 2002), Hurford (2012), and Progovac (this volume, chapter 5), among others. Bickerton's basic idea is that for a million or more years before modern language arose, hominids spoke something rather like today's pidgins. In fact, his notion of protolanguage lines up rather well with our notion of a concatenation grammar. Furthermore, the grammatical hierarchy offers a (relatively) graceful way for protolanguage itself to come into existence piecemeal, as well as for it to evolve piecemeal into fully complex language. In particular, the earlier stages don't go away—they are just elaborated upon and embedded in the more complex grammars.

Which of the levels of the hierarchy were actually instantiated in earlier hominids, and when? At the moment, we think there is no way to tell. Perhaps when more is understood about how brain structure creates specialized cognitive capacities, how genetics influences brain structure, and how the genome evolved from earliest hominids to modern humans, it will be possible to address these questions more rigorously. Meanwhile, we think the best that can be done is to keep one's eyes open for new evidence and try to develop theories of language that lend themselves to plausible evolutionary scenarios.

4.6 Conclusions

To sum up, we suggest that our grammatical hierarchy and the associated interface rules give us a useful way of pulling together a great number of disparate phenomena that fall under the radar of conventional generative grammar: issues in child language acquisition, late acquisition, language creation, language emergence, pidgins, language processing, language deficits, and the evolution of the language faculty.

A virtue of our analysis, we think, is that even fully complex languages display many symptoms of lower levels of the hierarchy, both in central domains such as word order and compounding and in more marginal constructions (see also Progovac, this volume, chapter 5). Moreover, there are full languages that appear to make little or no use of the upper reaches of the hierarchy. We think that our hypotheses lead to new and interesting conjectures about the nature of the human language faculty.

5

Degrees of complexity in syntax: a view from evolution

LJILJANA PROGOVAC

5.1 Setting the scene

Many syntacticians believe that syntax evolved suddenly, as a result of some minor, unremarkable event. Most recently, Berwick and Chomsky (2011: 29–31) assert that "the simplest assumption, hence the one we adopt . . . , is that the generative procedure emerged suddenly as the result of a minor mutation. In that case we would expect the generative procedure to be very simple . . . The generative process is optimal. . . . Language is something like a snowflake, assuming its particular form by virtue of laws of nature . . . Optimally, recursion can be reduced to Merge . . . There is no room in this picture for any precursors to language—say a language-like system with only short sentences. The same holds for language acquisition, despite appearances . . . " (for similar remarks, see several other chapters in Di Sciullo and Boeckx 2011). And according to Moro (2008: 53), there is no 'less-evolved' equivalent to the syntax of human languages.

The main goal of this chapter is to challenge this view of syntax and to argue that there is ample room for a grammatical system with only short (and flat) sentences, both in language evolution and in present-day languages.[1] My proposal is based on very specific claims, whose feasibility can be evaluated and tested both in the theory of syntax and in neuroimaging experiments. Using the formal framework of Minimalism, I argue that (root) small clauses, as well as intransitive absolutives (and

[1] For many good comments and discussions on this and related topics, I am grateful to, in no particular order, Eugenia Casielles, John L. Locke, Martha Ratliff, Jasmina Milićević, Ana Progovac, Draga Zec, Relja Vulanović, Steven Franks, Tecumseh Fitch, Fritz Newmeyer, Laurel Preston, Andrea Moro, Ray Jackendoff, Željko Bošković, Nataša Todorović, Igor Yanovich, David Gil, Dan Everett, Natasha Kondrashova, Dan Seeley, Richard Kayne, Juan Uriagereka, Rafaella Zanuttini, Stephanie Harves, Jim Hurford, Brady Clark, as well as many (other) audiences at various conferences. Special thanks go to the audience and organizers of the 2012 workshop on 'Formal Linguistics and the Measurement of Grammatical Complexity,' Seattle, Washington. All errors are mine.

middles), served as precursors to more complex language, given that they are syntactically simpler, and given that they provide a foundation for building more complex structures. There is evidence of evolutionary tinkering in the language design itself, and, as a result, language is neither perfect nor pure and hence in no sense resembles a snowflake. In other words, the complexity of syntax emerged gradually, through evolutionary tinkering, as suggested by, among others, Pinker and Bloom (1990); Newmeyer (1991); Jackendoff (1999, 2002); Culicover and Jackendoff (2005); Heine and Kuteva (2007); Hurford (2007); Givón (2009); Progovac (2006, 2009); and Jackendoff and Wittenberg, this volume, chapter 4.[2]

This chapter focuses on two concrete syntactic constructions: TP-less (root) small clauses, which are shown to be measurably simpler than finite TP counterparts (section 5.2); and intransitive absolutive (and unaccusative) clauses, shown to be measurably simpler than transitive counterparts (section 5.3). In this proposal, transitivity in syntax can be seen as an additional layer of structure, perhaps the vP layer of Minimalism (Chomsky 1995), superimposed upon the foundational (absolutive) layer. The simpler structures are argued to be evolutionary precursors to more complex counterparts, but both are still in use to varied degrees in present-day languages (section 5.4).[3] It is thus conceivable that languages vary with respect to how much they rely on the simpler, foundational structures, as opposed to the more recently emerged, and less robust, innovations.

The argument for each proposed progression through stages has three prongs to it: first, identifying 'living fossils' of each proto-stage in modern languages (in the sense of Jackendoff 1999, 2002), that is, measurably simpler syntactic constructions which nonetheless show continuity with more modern counterparts;[4] second, evidence of 'tinkering' with the language design, so that fossils of one stage intergrade into the next, leading to composite structures incorporating constructions of various stages; and third, existing or potential corroborating evidence from language acquisition,

[2] Pinker and Bloom (1990) propose that language evolved gradually, subject to the Baldwin Effect, the process whereby environmentally-induced responses set up selection pressures for such responses to become innate, triggering conventional Darwinian evolution (see also Deacon 1997; Hinton and Nowlan 1987). Tiny selective advantages are sufficient for evolutionary change; according to Haldane (1927), a variant that produces on average 1% more offspring than its alternative allele would increase in frequency from 0.1% to 99.9% of the population in just over 4,000 generations. This would still leave plenty of time for language to have evolved: 3.5–5 million years, if early Australopithecines were the first talkers, or, as an absolute minimum, several hundred thousand years (Stringer and Andrews 1988), in the event that early *Homo sapiens* was the first. Moreover, fixations of different genes can go in parallel, and sexual selection can significantly speed up any of these processes (for possible evidence of sexual selection for (simple) syntax, see Progovac and Locke 2009).

[3] This evolutionary scenario can shed light on the availability of a variety of linguistic strategies for marking case in present-day languages, as well as on the availability of certain intermediate 'middle' constructions, which may have facilitated transition toward transitivity.

[4] In the biological literature, living fossils are defined as species that have changed little from their fossil ancestors in the distant past, e.g. lungfish (Ridley 1993). Linguistic fossils are also discussed in Bickerton (1990, 1998).

loss, neuroimaging, and genetics (section 5.5). Section 5.5 also points to the abundant literature on processing which shows that increased syntactic complexity leads to increased activation in certain areas of the brain.

The recurring theme of this work is that each stage preserves, and builds upon, the achievements of the previous stage(s). Thus, a TP is built upon the foundation of the small clause. Likewise, transitive structures (vP shells), as well as 'middles,' are built upon the foundation of intransitive (absolutive-like) VPs (or small clauses (SCs)). Several other works share an approach which assumes the layering and preservation of older stages, such as Jackendoff and Wittenberg (this volume, chapter 4). In brain stratification accounts, such as are found in the work of Vygotsky and Piaget, as well as in proposals for a triune brain (MacLean 1949), the common theme is the inclusion of the attainments of earlier stages in the structures of later stages. According to Vygotsky (1979/1960: 155–156), "instinct is not destroyed, but 'copied' in conditioned reflexes as a function of the ancient brain, which is now to be found in the new one."

My claim is thus that modern languages make use of constructions of unequal (morpho-) syntactic complexity to express the same or similar meanings. First, as pointed out above, languages can vary with respect to how much they rely on foundational constructions, as opposed to grammatically more layered structures. Second, even internal to a single language, one can find alternative ways to express similar meanings. I argue that acknowledging and characterizing variability in syntax is a necessary condition, not only for exploring the evolutionary origins of language, but also for formulating useful hypotheses regarding how syntax is represented in the brain, how it is acquired, and how it is affected in language disorders. All of this puts us in a better position to interpret, as well as to propel, new findings in genetics and neuroscience.

5.2 (Root) small clauses as evolutionary precursors to Tense Phrases (TPs)

Several researchers have proposed that there was a paratactic stage of syntax in the evolution of human language, involving two elements loosely concatenated into a single utterance (see Givón 1979; Dwyer 1986; Bickerton 1990; Jackendoff 1999, 2002; Culicover and Jackendoff 2005; Jackendoff and Wittenberg, this volume, chapter 4; Deutscher 2005; Burling 2005). Using data from English and Serbian, Progovac (2006, 2008a,b, 2009) has applied this idea to what are referred to in the literature as 'small clauses,' which are found both integrated into more complex sentences and in isolation as root clauses. According to Progovac, root small clauses, illustrated in (1), lack the TP (Tense Phrase) layer of structure typically associated with modern sentences in Minimalism, and are thus grammatically measurably simpler than the full TP counterparts illustrated in (2):

(1) Tim happy? Him worry? Problem solved. Case closed. Point taken. Mission accomplished. Crisis averted. Family first! Me first! Everybody out!
(2) Tim is happy. He worries? The problem has been solved. The point is taken. I will/should be first!

Serbian has (at least) two ways of expressing propositions with unaccusative verbs: first, as VS (Verb-Subject) small clauses arguably without the TP layer (3a and 4a) and second, as full finite TPs with free order (3b, 4b) (for details, see Progovac 2008a,b):[5]

(3) a. Stigla pošta.
arrived.PP mail

b. Pošta je stigla. / Stigla je pošta.
mail AUX arrived
[PP stands for past participle]

(4) a. Pala karta. / a'. ?*Karta pala.
fell. PP card
'Card laid, card played.'

b. Karta je pala. / Pala je karta.
'The card fell.'

In addition to the obvious morphosyntactic hallmarks of root small clauses, such as the absence of a finite verbal form, absence of nominative case checking in English, and absence of subject raising in Serbian (hence rigid VS order (4a–4a')), these clauses are also characterized by the following surprising properties: they do not tolerate movement of any kind (5–6); they cannot embed one within another, and thus do not show recursion (7–9); their interpretation is typically confined to the here and now (10–11); and many among them are (semi-)formulaic.[6] For all these reasons, these clauses cannot be analyzed as identical in structure or complexity to their full finite counterparts, nor can they be analyzed as elliptical versions of the full counterparts:

[5] Roughly speaking, unaccusatives are intransitive structures whose subject is not an agent, but rather a theme/patient, and thus object-like in certain ways (see Perlmutter 1978 and Burzio 1981 for early characterizations and analyses).

[6] Move and recursion are considered to be universal and defining properties of human language among Minimalist researchers (see e.g. Hauser, Chomsky, and Fitch 2002; Chomsky 2005; Fitch, Hauser, and Chomsky 2005; Moro 2008). If considerations in this chapter are on the right track, then Move and recursion cannot be the defining properties of human language, but rather relatively recent, fragile innovations. See below in the text for the claim that Proto-Indo-European also lacked clausal embedding and recursion. Newmeyer (2005: 170–171) leaves the door open for languages to lack subordination, suggesting that such a lack may be correlated with the lack of literacy (see references cited there).

(5) *Who(m) worry?!
*Where everybody?!

(6) *Kada stigla pošta? (cf. Kada je stigla pošta?)
when arrived mail

(7) *Him worry [me first]?

(8) *Him worry [problem solved]?

(9) *Ja mislim [(da) stigla pošta]. (cf. Ja mislim [da je stigla pošta].)
I think (that) arrived mail

(10) *Stigla pošta pre tri godine.
three years ago
*Pala karta pre tri godine.

(11) *Case closed three years ago.
*Me first three years ago!

I conclude that two distinct types of grammar are at play here: the simpler, rigid, (TP-less) small clause (SC) grammar, approximating the ancient paratactic stage, and the more complex TP grammar, which subsumes the former in that a TP is built upon the SC foundation. TP clauses have at least one more layer of structure than root small clauses (or 'half-clauses,' as they are referred to in Progovac 2008b).[7] Superimposing one layer (e.g. TP) over another (SC) creates hierarchy, as well as additional syntactic space for Move to target as it connects multiple layers of structure. Clearly, the reason why clauses in (7–9) cannot embed has nothing to do with the speakers not being capable of recursive thought, and everything to do with the structure of these clauses.

This kind of derivation (from SC to TP) is the commonly accepted postulate in Minimalism and predecessors, dating back to Stowell (1981, 1983); Kitagawa (1986); Koopman and Sportiche (1991) (see also Burzio 1981; Bošković 1995; Hale and Keyser 2002; Chomsky 1995; and subsequent Minimalist work). In fact, the derivation of sentences from small clauses is one of the most stable postulates in the framework, having survived many changes of analysis and focus. In other words, as the examples in (12–13) illustrate, small clauses morph/transform into TPs, as if the building of the modern sentence (TP) retraces its evolutionary steps.

(12) a. Small Clause: [$_{SC}$ Sheila sad] →
b. [$_{TP}$ is [$_{SC}$ Sheila sad]] →
c. [$_{TP}$ Sheila [$_{T'}$ is [$_{SC}$ ~~Sheila~~ sad]]]

[7] As put in Carroll (2005: 170–171), "the erroneous notion . . . has been that the intermediate stages in the evolution of structures must be useless—the old saw of 'What use is half a leg or half an eye?'" My argument is that these 'half-clauses' would have been immensely useful to our ancestors, especially at the point when they first broke into language.

(13) a. Small clause: [$_{\text{SC}}$ pala vlada] →
fell government

b. [$_{\text{TP}}$ je [$_{\text{SC}}$ pala vlada]]→
has fallen government

c. [$_{\text{TP}}$ vlada [$_{\text{T'}}$ je [$_{\text{SC}}$ pala ~~vlada~~]]]
'The government has collapsed.'

Due to the wiring of the brain to accommodate this kind of sentence derivation throughout the evolution of human language, it is possible that the only way that humans can build sentences is by relying on the small clause, even though one can envision more direct and more optimal derivations.

The TP grammar allows for embedded recursion (14) and for the expression of a variety of nuanced meanings with respect to the temporal/aspectual properties of the clause (15). The small clause grammar, on the other hand, allows for flat concatenation of the type illustrated in (16–17), once again often resulting in (semi-)formulaic expressions that resist questioning and recursion (for details, see Progovac 2010a):

(14) He worries [that I think [that the problem has been solved]].

(15) The problem has been/may have been/will be solved.
I will be/should be/better be first.

(16) Nothing ventured, nothing gained.
Easy come, easy go.
Monkey see, monkey do.

(17) Preko preče, naokolo bliže. (Serbian)
'Across shorter, around closer.'

The discussion in this section shows that Minimalism is equipped with formal tools which can characterize constructions of varied degrees of syntactic complexity. Not only that, but the theoretical postulate of the multiple layers of clausal structure provides a particularly suitable platform for this task. Given this approach, it is conceivable that some languages make predominant or sole use of small clause grammars, as suggested by the descriptions of, for example, Riau Indonesian in Gil (2005a) and Pirahã in Everett (2005). Indeed, Kiparsky (1968) has argued that Proto-Indo-European (PIE) (spoken around 3700 BC) had optional adverbial temporal particles, which did not build TPs, but should rather be analyzed as adjuncts.[8]

[8] According to Kuryłowicz (1964: 21), the injunctive, a tenseless verbal form, was the only mood in earliest PIE. Gonda (1956: 36–37) points out that any attempt exactly to translate the injunctive categories into a modern Western idiom is doomed to fail, given "the vagueness in meaning and the great, and in the eyes of modern man astonishing, variety of its functions."

Adjunction itself, used abundantly in modern languages, can be seen as a living fossil of the paratactic proto-grammar (see Jackendoff 1999, 2002).

In addition, according to Kiparsky (1995: 155), a major characteristic of PIE syntax, best preserved in Sanskrit, Hittite, and Old Latin, was that finite 'subordinate' clauses were not embedded but rather (paratactically) adjoined (see also Hale 1987). According to Kiparsky, PIE lacked the category of Complementizer and had no Complementizer Phrase (CP) or any syntactically embedded sentences. If so, then PIE clauses resemble root small clauses.

One piece of corroborating evidence for a small clause stage in language history comes from language acquisition: it has been argued by many that children go through a small clause/root infinitive stage (Radford 1988, 1990; Lebeaux 1989; Platzak 1990; Ouhalla 1991; Rizzi 1994; Potts and Roeper 2006; but see Guasti 2002 for opposing views).[9] When it comes to aphasia, Kolk (2006; also references there) has argued that there is bias in agrammatic speakers of Dutch and German to select simpler types of constructions, often small clauses/root infinitives (see also Friedmann and Grodzinsky 1997). Whereas control speakers in Kolk and colleagues' studies produced 10% nonfinite root clauses, aphasics produced about 60%.[10] Last but not least, the hypothesis of the small clause stage in the evolution of syntax is both testable and falsifiable: one can use the subtractive neurolinguistic method to test how TP-less small clauses are processed in comparison to the more complex TP counterparts, as discussed in more detail in section 5.5 (see also Progovac 2010b).

Identifying clauses of differing degrees of syntactic complexity makes it possible to ask meaningful questions about how syntax evolved, how it is acquired, as well as how it is represented in the brain or affected in language disorders. As the next section shows, this approach can also shed light on various puzzling properties of language design, including unaccusativity and ergativity.

5.3 Intransitive absolutives (and middles) as evolutionary precursors

Notice that the examples of small clauses considered in the previous section are all intransitive, consisting of just two words, a predicate and one argument. I chose such examples on purpose as it is quite conceivable that the first combinations of (proto-) words were two-word utterances. A two-word stage has also been postulated in first language acquisition (e.g. Bloom 1970).

[9] Even though ontogeny (development in children) does not literally recapitulate phylogeny (development in species), according to Studdert-Kennedy (1991), Rolfe (1996), Locke (2009) and others, present-day views warrant the use of ontogeny to corroborate hypotheses about phylogeny. In biological texts (e.g. Ridley 1993), the relationship between ontogeny and phylogeny is considered to be a classic topic in evolutionary studies.

[10] According to Kolk, the overuse of nonfinite root clauses by children decreases with age: from 83% in the 2-year-olds, to 60% in the 2.5-year-olds, to 40% in the 3-year-olds.

Consistent with these considerations, this section postulates a stage in the evolution of syntax in which only intransitive absolutive-like patterns were available, i.e. patterns in which a verb takes only one argument, and in which the distinction between subjects and objects is neutralized. In fact, that kind of a distinction would be irrelevant in a grammar which licenses only one argument per clause. Such a view is very different from one that would assume missing or null arguments. In my view, transitivity is an evolutionary innovation, that is, an additional layer of structure (perhaps the vP layer of the Minimalist framework) superimposed upon the foundational absolutive (small clause) layer. Pressure to accommodate additional arguments would have been a driving force behind the evolution of more complex (transitive) patterns, with some middle and *se*-type constructions paving the way toward transitivity. The living fossils of this stage would include exocentric VN compounds, unaccusatives, absolutives, as well as (other) absolutive-like constructions found in nominative-accusative languages, as will be discussed below. This approach can shed light on the existence of different case marking patterns in the world's languages, including ergative-absolutive and nominative-accusative, as well as the availability of the foundational absolutive-like pattern in various guises in primarily nominative/accusative languages.

As transitivity evolved, that is, in the process of grammaticalization of the syntactic positions of more than one argument, there are/were, in my view, various types of intermediate steps. The fossils of these intermediate stages include so-called 'middle' constructions, which straddle the boundary between transitivity and intransitivity, as well as between passives and actives, and which neutralize the distinction between subjects and objects. I exemplify such middle grammars with *se* constructions in Serbian, where *se*, rather than being a reflexive, is best analyzed as an expletive pronoun superimposed over an absolutive SC layer. Such middle constructions are more common in some languages than in others.

The grammaticalization of transitivity in nominative/accusative languages, which have a structural accusative case and a vP/VP shell, would not have precluded some other structures (e.g. unaccusatives, *se* clauses, nominals, compounds) from retaining an absolutive flavor. If grammars containing these simpler structures involve less syntactic processing, then their retention at least in some constructions is beneficial.

There is some corroborating evidence for an intransitive absolutive stage in the evolution of human language. I discuss it at this point because it can help one envision not only what such a stage would have looked like, but also why there would be pressure to develop transitivity. Consider the emergence of Nicaraguan Sign Language (NSL) among deaf children in the 1970s and 1980s. According to Kegl *et al.* (1999: 216–217), the earliest (pidgin) stages of NSL did not use transitive NP V NP constructions with animate nouns, such as (18). Instead, an alternative strategy was used, consisting of two paratactically combined (intransitive) clauses, an NP V — NP V sequence (19–20):

(18) *WOMAN PUSH MAN.
(19) WOMAN PUSH — MAN REACT.
(20) WOMAN PUSH — MAN FALL.

Similarly, Goldin-Meadow (2005) reports that in homesign syntax, patients/themes tend to precede verbs, as do intransitive actors, resembling an ergative system (e.g. BOY-HIT, APPLE-EAT). In addition, patients are more likely to be expressed than agents. Given that these sign languages are created from scratch, they can provide a window into the beginnings of human language.

5.3.1 *An intransitive (absolutive) stage: what unaccusatives, exocentrics and absolutives have in common*

In Minimalism, transitivity is considered to involve an additional layer of verb structure, a vP/VP shell (e.g. Chomsky 1995), where the internal (closer argument) is generated in the VP (or SC), and the external argument (e.g. agent) in the vP (see 21–22). Intransitives, especially, unaccusatives, can be accommodated without the vP layer (23–24):[11]

(21) Maria will roll the ball.

(22) a. [$_{SC/VP}$ roll the ball] →
b. [$_{vP}$ Maria [$_{SC/VP}$ roll the ball]] →
c. [TP: Maria will [$_{vP}$ ~~Maria~~ [$_{SC/VP}$ roll the ball]]]

(23) The ball will roll.

(24) a. [$_{SC/VP}$ roll the ball]
b. [TP: The ball will [$_{SC/VP}$ roll ~~the ball~~]]

As shown in (23–24), intransitives can be derived without a vP layer. If there is no TP layer either, then they can remain bare small clauses, as in the examples of unaccusative root small clauses introduced in the previous section, one of which is repeated here:

(25) [$_{sc}$ pala vlada] (Serbian)
fell government

The unaccusativity phenomenon can thus be seen as an option to retain (elements of) proto-grammars in constructions which can be supported by such grammars, i.e. intransitive constructions with a single internal argument. Notice that unaccusatives blur the subject/object distinction: their only argument has properties of both.

[11] According to Borer's (1994) fully configurational approach to argument linking, the arguments within the VP are hierarchically unordered and there is no lexical distinction between subjects and objects in the VP. Such distinctions can only be made with the help of functional projections such as vP.

Another phenomenon that is difficult to explain given the postulates of modern morphosyntax is exocentric VN compounds of the kind illustrated in (26–27) below. According to Progovac (2009, 2012a) and Progovac and Locke (2009), they are fossils closely approximating an absolutive (intransitive) stage in the evolution of human language.

(26) scare-crow, kill-joy, pick-pocket, cry-baby, spoil-sport, turn-coat, rattle-snake, hunch-back, dare-devil, wag-tail, tattle-tale, pinch-penny (miser), sink-hole, busy-body

(27) ispi-čutura (drink.up-flask—drunkard), guli-koža (peel-skin—who rips you off), cepi-dlaka (split-hair—who splits hairs), muti-voda (muddy-water—trouble-maker), jebi-vetar (screw-wind—charlatan), vrti-guz (spin-butt—fidget); tuži-baba (whine-old.woman; tattletale); pali-drvce (ignite-stick, matches) (Serbian)

The only argument in these compounds is often internal, corresponding to a sentential object, but it can also be external, corresponding to a sentential subject, as in the underlined compounds. The grammar behind these compounds provides no morphosyntactic differentiation between internal and external arguments, resulting in vagueness. For example, a *rattle-snake* is conventionally interpreted as a snake that rattles, but one can imagine this word also used for somebody who routinely rattles snakes. Notice that a more complex compound, *snake-rattler*, which has a transitivity layer, can only be interpreted as somebody who rattles snakes. I conclude that these compounds exhibit only the rudiments of predication (call it proto-predication), where the only argument is assigned a thematic role by the verb, but the nature of that thematic role is left unspecified.

This is consistent with the claim in Gil (2012b, this volume, chapter 3) that predication is a composite emergent entity, rather than a primitive, and that it brings together both thematic role assignment and headedness. Casielles and Progovac (2010, 2012) relate proto-structures involving unaccusatives to thetic statements, which have also been argued by philosophers and linguists to be pre-predication structures.

The vagueness attested in these compounds is also characteristic of absolutives in some ergative-absolutive languages. Consider the following example from Tongan featuring a typical intransitive sentence with absolutive case (Tchekhoff 1979: 409):[12]

[12] See also Gil (2005a; this volume, chapter 3) for an extensive discussion of comparable vague clauses in Riau Indonesian:

(i) Ayam makan
chicken eat
'The chicken is eating.'
'Somebody is eating the chicken.'

(28) ‘oku kai ‘ae iká.
PRES eat the fish
‘The fish eats.’
‘The fish is eaten.’

In ergative-absolutive patterns, the subject of an intransitive predicate is characterized as morphosyntactically equivalent to the object of a transitive predicate, both occurring in the absolutive case (e.g. Comrie 1978; Dixon 1994). It is only after the addition of an external (agent) argument (e.g. *the man*), presumably in the vP layer (but see Alexiadou 2001 and Nash 1996, for an adjunction analysis of the ergative argument),[13] that the role of *the fish* disambiguates and is necessarily interpreted as patient/theme (the addition of –*er* in verbal compounds has a comparable effect, as pointed out above). In other words, the addition of an outer layer forces the inner layer to reorganize and specialize with respect to the outer layer, resulting in more precision. Assuming that there was an intransitive absolutive (proto-syntactic) stage in the evolution of human language, one can envision the subsequent development of two language types, primarily nominative-accusative and primarily ergative-absolutive.[14]

Bringing unaccusativity and ergativity under the same umbrella, Bok-Bennema (1991: 169) argues that ergativity and unaccusativity are both characterized by the inability of transitive verbs to assign structural case. According to e.g. Alexiadou (2001: 18; also Hale 1970; Nash 1995), ergative/absolutive patterns are reflexes of a passive/unaccusative system. Therefore, what absolutives, exocentric VN compounds, and unaccusatives have in common is that the verb is unable to assign structural case to its argument, resulting in the blurring of the distinction between subjecthood and objecthood. These phenomena begin to make sense if they are seen as approximations of a proto-syntax stage, which could accommodate only one argument per clause.

As pointed out above for unaccusatives, the absolutive patterns might have been preserved in some constructions because they are structurally simpler and thus involve less syntactic processing. It is also conceivable under this approach that the foundational absolutive-like patterns will be found in some guise or another in nominative/accusative languages as well, as explored in the following section. Needless to say, languages may vary considerably with respect to the degree to which they rely on the foundational absolutive patterns.

[13] According to Nash (1995: 119), only nominative/accusative languages have external arguments in vP; ergative languages have their agents internally within the VP, possibly as adjuncts.

[14] Lehmann (1985a: 245) points to the gradient nature of the distinction between the ergative, active, and accusative types: “a language is never wholly and exclusively either ergative or active or accusative, in all its grammatical patterns.”

5.3.2 *Absolutive patterns in nominative/accusative languages*

The previous section concludes that the only argument in exocentric VN compounds is absolutive-like. In addition, Alexiadou (2001) argues that nominals across various languages are intransitive, as well as absolutive-like (passive-like). In other words, all nominals, whether passive or not, have an intransitive base (see also Picallo 1991; Bottari 1992; Alexiadou and Stavrou 1998). In passive nominals the agent appears as an adjunct, as in (29) from Alexiadou (2001: 78).[15]

(29) the destruction of the city <u>by the barbarians</u>

By-phrases in derived nominals can be interpreted only as affectors (agents, instruments, creators), which explains the ungrammaticality of sentences such as (30).

(30) ?*the receipt of the package by John

Unlike with the verbal domain, there is no structural external argument in nominalizations, and the presence of the *by*-phrase seems to be lexically licensed, as suggested by (30). In that sense, the external argument in the *by*-phrase resembles the ergative argument in ergative languages which is also often analyzed as surfacing in lexical/prepositional case, rather than structural case, as we saw in the previous section.

The following examples from Serbian also illustrate the intransitive nature of nominals. (31) shows that only one argument can get structural case in the noun phrase, while (32) shows that the agent can be introduced in an oblique phrase also found in passives.

(31) kritikovanje studenata (*profesora) (Serbian)
criticizing students.$_{\text{GEN}}$ professors.$_{\text{GEN}}$
'criticizing of the students (*of the professors)'

(32) kritikovanje studenata od strane profesora
criticizing students.$_{\text{GEN}}$ by professors.$_{\text{GEN}}$

Consider next dative 'subjects' with *se* in Serbian, which co-occur with nominative 'objects' in what looks like an ergative/absolutive pattern:

(33) Meni se pije kafa. (Serbian)
me.$_{\text{DAT}}$ SE drinks coffee.$_{\text{NOM}}$
'I feel like drinking coffee.'

[15] Comrie (1978) suggests that nominalizations constitute a possible source for ergativity. Or perhaps it is the other way around.

Nominative marking on the 'object' is like absolutive, being also the case of intransitive subjects, while dative introduces an external argument, akin to an ergative (see e.g. Alexiadou 2001; Nash 1996, for an adjunction analysis of the ergative argument). According to Nash (1996: 171), ergative subjects, like dative subjects, cannot co-occur with structural accusative, but instead appear with absolutive/nominative 'objects.' Dative subject clauses are yet another construction in which the verb is unable to assign structural (accusative case) to what would be its object.

It is also of significance here that dative subjects in Serbian co-occur with the pronoun *se*. As we see in the next section, *se* is associated with the ancient absolutive pattern, thereby 'modernizing' this pattern to some extent by providing a transition toward transitivity.

5.3.3 *The 'middle' ground*

According to e.g. Kemmer (1994: 181), "the reflexive and the middle can be situated as semantic categories intermediate in transitivity between one-participant and two-participant events." Let us consider some *se* constructions in Serbian that can be characterized as middle constructions or as constructions straddling the boundary between the passive and active voice.

In addition to dative subject clauses introduced in the previous section, *se* also occurs in Serbian equivalents of the so-called ergative verbs (verbs which exhibit both transitive and intransitive patterns, as in the English example (34) from Radford 2004). *Se* only occurs, and obligatorily so, when such verbs are used intransitively, as in (35):[16]

(34) The ashtray broke. / He broke the ashtray.

(35) a. Pepeljara se razbila.
ashtray SE broke

b. On je razbio pepeljaru.
he AUX broke ashtray

Where pragmatics allows, *se* constructions in Serbian exhibit astonishing vagueness of meaning:

(36) Deca se tuku.
children SE hit
'The children are hitting each other/themselves.'
' The children are hitting somebody else.'
'One hits/spanks children.'

[16] In a similar fashion, according to Arce-Arenales, Axelrod, and Fox (1994), some Spanish transitive verbs can be 'made intransitive' with the use of *se*.

(37) Pas se ujeda.
dog SE bites
'The dog bites (someone).'
? 'The dog is biting itself.'
? 'One bites dogs.'

(38) Marko se udara.
Marko SE hits
'Marko is hitting **me**.'
'Marko is hitting somebody.'
'Marko is hitting himself.'

If (38) is uttered with a sense of urgency, the most probable interpretation will involve the most salient discourse participant, the speaker, even though there is no morpheme corresponding to 1st person at all! My argument is that *se* simply implies that there is one more participant involved in the event, in addition to the one surfacing, and typically its role can be inferred from pragmatic context (e.g. 38). But the role of the expressed argument (e.g. *Deca*, *Pas*, or *Marko* above) still remains absolutive-like, not grammatically specified as either subject or object (Progovac 2005, 2012b).

Comparable vagueness is also found with *se* constructions in Spanish (Arce-Arenales, Axelrod, and Fox 1994: 5):[17]

(39) Juan se mató.
Juan SE killed
'Juan got killed.'
'Juan killed himself.'

Serbian *se* is analyzed in Franks (1995) and Progovac (2005) as an expletive pronoun, 'absorbing' accusative case. Notice that the vagueness in *se* clauses can be compared to that found in Tongan (28) and Riau intransitives (see footnote 12).

My proposal is that *se* constructions (and middles in general) involve less syntactic processing than regular transitives.[18] As suggested in Progovac (2012b), this is something that can be tested given, for example, the minimal contrasts in Serbian between *se* constructions (40) and true transitive counterparts featuring the 1st person pronoun *me* (41) (see section 5.5):

[17] Compare also the vagueness/ambiguity of the English example below:

(i) The children got dressed.

[18] My proposal that *se* constructions are syntactically simpler than transitives is consistent with the finding that reflexive constructions are easier for agrammatic patients to process than true transitive constructions (e.g. Grodzinsky *et al.* 1993).

(40) Marko se udara.
(41) Marko me udara.

Se can thus be seen as (gradually) introducing (grammaticalized) transitivity/accusative marking to the ancient absolutive pattern. Notice that without *se*, the absolutive pattern vanishes, and the only argument has to be interpreted as subject/agent performing an action on an unspecified object, as is also the case with English translations:

(42) Deca tuku.
'The children are hitting (somebody).'
(43) Pas ujeda.
'The dog bites (someone).'

Se constructions in Serbian are very commonly used; in addition to appearing instead of true transitives, as per the discussion above, they are also used with ergative verbs, as illustrated in (35), as well as where English speakers would use a passive.

It seems, then, that the distinctions between subjecthood and objecthood, transitivity and intransitivity, and passive and active, can be neutralized, and can have a middle ground. One way to make sense out of this is to postulate an intransitive absolutive-like stage in the evolution of human language, which provides a foundation for any subsequent elaboration of argument structure. Importantly, however, introducing transitivity with a structural accusative case (vP/VP shell) does not preclude some other constructions (e.g. unaccusative small clauses, nominals, *se*-constructions, compounds) from remaining absolutive-like.

Reinforcing the evolutionary argument, many of these foundational structures still live inside the more complex structures: absolutives arguably live inside nominals, *se* constructions, and transitives, while small clauses live inside TPs. In other words, small clauses and intransitive absolutives constitute the foundation, the platform on top of which one can build (or not) more complex syntax, namely TPs and transitivity. This is not to say, however, that evolution can never revert back to the more robust, foundational strategies. According to the so-called 'last in, first out' principle, used in computer science and psychology (see Code 2005), what is acquired last is the most shallow layer that is the easiest to lose, and vice versa. When it comes to complex syntax, the loss can take place in pidginization and in agrammatic aphasia (see also Gil 2005a for the development of Riau Indonesian).

While we often think of biological evolution as a linear progression toward higher complexity, some recent genetic studies reveal that reversals and losses are certainly possible even in the evolution of some truly important traits, such as multicellularity, a major transition in the history of life.[19] For example, a recent article in *BMC*

[19] A much more banal example would be the possibility for animals to acquire fur and then to lose it.

Evolutionary Biology by Schirrmeister, Antonelli, and Bagheri (2011) reports that the majority of extant cyanobacteria, one of the oldest phyla still alive ('living fossils'), including many single-celled species, descend from multicellular ancestors, and that reversals to unicellularity occurred at least five times. According to their findings, most of the morphological diversity exhibited in cyanobacteria today—including the majority of single-celled species—arose from ancient multicellular lineages.[20] In a sense, then, pidginization can be seen as equivalent to the return to a simpler, unicellular mode of existence.

Finally, as with the small clause/TP distinction, the hypotheses explored in this section are testable/falsifiable. One can use the subtractive neurolinguistic method to test how syntactically simpler structures, such as absolutives and middles, are processed in contrast to the more complex counterparts (section 5.5).

5.4 What creates complexity in syntax, and why it matters

Complexity grows when more tools and operations become available, with new and old tools starting to specialize with respect to each other's functions, similar to the way that vocabulary items specialize when new words are introduced. For example, the borrowing of the word *beef* caused the word *cow* to specialize for everything but table meat, which is the territory covered by *beef*. No such bleeding occurs with the words *frog* or *chicken*, where equivalent borrowings did not take place. Availability of more words in a particular domain leads to more specialization and precision, but at the cost of losing the ability to use the old words in their more vague, original senses. All other things being equal, I don't think that anybody would dispute that the lexicon involving a *beef/cow* distinction is more complex than the lexicon which has only the word *cow*.

By the same reasoning, a system that utilizes both (root) small clauses and TPs is more complex than a system that only utilizes small clauses. As Carroll (2005: 170–171) puts it, "multifunctionality and redundancy create the opportunity for the evolution of specialization through the division of labor…" It is of note here that small clauses and TPs in languages with both exhibit a division of labor, at least to some extent. TPs are used in root contexts, while small clauses are typically confined to being constituents within full sentences. When used in root contexts, small clauses, in contrast to TPs, tend to express utterances in the here-and-now, as well as formulaic and semi-formulaic utterances, as established in section 5.2. This situation is not unlike the one involving additional vocabulary items.

[20] The authors also mention that a complex character such as eusociality has been lost several times in halictid bees (see e.g. Danforth, Conway, and Shuqing 2003). Eusociality refers to the highest level of social organization in a hierarchical classification.

Moreover, the more functional projections such as TP, CP, and DP get projected on top of a lexical projection, the more complex the output. For this reason, a transitive structure with a vP/VP shell is more complex than an intransitive structure with only a bare VP (or SC), and a system which uses both bare VP clauses and vP shell clauses is more complex than a system which uses only the former.[21]

Linguists often claim that all languages are equally complex because one can express any thought in any language. Even if that is true, it does not follow that all languages have equally complex grammars. After all, one can communicate a thought even without speaking at all, by, for example, pointing to a tree. Perhaps lightning hit the tree and the pointing draws attention to that fact, or perhaps it warns the hearer to stay away from the tree. Pointing is a vague act of communication, heavily dependent on the context, but so are many expressions in language, including one-word utterances, such as *Tree*! In fact, any expression in any language is vague along some dimension, vagueness being a matter only of degree.

In other words, while one can indeed express many different thoughts using language, or even without using language, that does not mean that all languages avail themselves of an equally complex inventory of grammatical tools. A good chef can prepare a fantastic meal by using a single all-purpose knife, but that does not mean that his or her set of tools is equally complex and elaborate as that of another chef, who uses ten different specialized knives. Making a claim that one set of tools is more elaborate or complex than another does not imply that the chef with ten knives will prepare a more satisfying meal, nor that speaking in TPs, DPs and vP shells will bring more wisdom to the speakers.[22] It has been argued by certain minimalists (Bošković 2008, for example) that some languages, like English and French, project a DP projection above the NP, while others, like Japanese and Serbian, do not. This is still a controversial matter, but if the insight is on the right track, this is not taken to mean that English or French speakers are superior to Japanese or Serbian speakers, although it could mean that their brains process noun phrases differently. That would be a really interesting and testable question for research. If it turns out to be true, then English has an additional grammatical tool to allow the automaticization of properties like (in)definiteness, and perhaps to make it possible to express possessive relationships recursively (e.g. *John's mother's best friend's bike*). In Serbian, for

[21] In the same vein, a grammar which has Merge but not Move is simpler than a grammar which has both (Progovac 2009), although the lack of Move may be reducible to the lack of hierarchical structure, which characterizes paratactic grammars. As established in section 5.2, small clause grammars do not exhibit Move.

[22] At the same time, a chef with ten specialized knives is probably in a better position to decorate the meal with fancy rose-petal garnishes, as pointed out by Laurel Preston, p.c. Perhaps the recursion in the noun phrase, as discussed below, can be seen as such a fanciful decorative touch in the realm of syntax. Or perhaps recursion in the noun phrase is an epiphenomenon of another evolutionary development, DP. While there are of course no immediate answers to these questions, at least they can be raised and addressed in this evolutionary framework.

example, this kind of recursion in the NP is impossible; in addition, possibly also due to the lack of DP in their native language, speakers of Serbian have a hard time learning how and when to use articles in English.

Linguists often also claim that if one language has less complexity in one area of grammar, then it compensates for it in another (see e.g. Newmeyer and Preston, this volume, chapter 1, for references and discussion). It is not clear by what grammatical means Serbian speakers might compensate for the lack of recursion in the noun phrase. As Gil (this volume, chapter 3) points out, there is no reason to believe "that speakers of a language such as Riau Indonesian always worry about whether a clause unmarked for TAM [Tense-Aspect-Mood] has past, present or future reference any more than speakers of a language such as English invariably puzzle over past-tense-marked clauses, wondering whether they refer to a few hours back, a day ago, a week to a month ago, a couple of months to a couple of years ago, or the distant or legendary past, just because these distinctions happen to be grammaticalized in a language such as Yagua." In other words, English does not have to compensate for the lack of distinctions that Yagua makes, nor does Serbian need to compensate for the lack of distinctions that English makes in the DP. It is not clear at all what area of the brain, or what other entity, would be responsible for monitoring the overall complexity of each language, making sure that it has neither more nor less than any other language in the world. It is even less clear why such an entity should exist.

The interesting question is how the brain processes syntactic constructions of unequal grammatical complexity. It should also be of interest to linguists to determine if grammatical structure helped shape other aspects of our brains during human evolution. While linguists often assert that language is somehow a (passive) consequence of an evolved brain, it may well be that the pressure to grammaticalize and make automatic various aspects of language directly contributed to the increase in the size of the brain, as well as to its lateralization. If linguists do not put language in the center of these investigations, nobody else will. The only way to explore the question of how language and the brain (co-)evolved is to explore specific hypotheses about language evolution, recognizing variability in complexity, not only across time, but also across languages and constructions.

5.5 Testing grounds for variability in syntactic complexity

As pointed out throughout the chapter, neuroimaging can provide a fertile testing ground for various evolutionary claims, including the hypotheses explored here. Conversely, the advancement and testing of specific hypotheses about language complexity in relation to the evolutionary origins of language puts us in a better position to interpret, as well as to spearhead, new findings in genetics and neuroscience.

According to Poeppel and Embick (2005), among others, current neurolinguistic research presents a case of cross-sterilization, rather than cross-fertilization. This is because, according to them, no meaningful correlates have been found, nor are expected to be found, between biological units of neuroscience, such as neurons, dendrites, axons, and constructs from formal syntax such as Move, Subjacency, and the Theta-Criterion. In my view, this gridlock can be broken by means of the subtractive neurolinguistic method, widely used in neuroscience. Very roughly speaking, this method compares and contrasts how certain constructions are processed in the brain by subtracting the brain image resulting from the processing of one construction from that of another, isolating the differences between the two.

The method should allow us to determine how proto-syntactic structures such as root small clauses, absolutives, and middles are processed in comparison to their more complex counterparts. The result, one hopes, will be the discovery of neurobiological correlates of, for example, TP-layering and vP shells (see Progovac 2010b). In this concrete case, the subjects can be exposed to sentences such as *Marko me udara* and *Marko se udara*; *Me first!* and *I am first*! The images for such minimal pairs would then be subtracted from each other. The differences in images between *Marko me udara* and *Marko se udara* are expected to isolate the brain activity associated with true transitivity (e.g. the projection of the vP shell in Minimalism) (section 5.3.3), while the differences in images between *I am first!* and *Me first!* should isolate the brain activity associated with the TP layer (section 5.2).

The processing of TPs and transitives with vP shells should show clear lateralization in the left hemisphere, with extensive activation of the Broca's area. On the other hand, proto-structures, such as root small clauses and absolutive-type constructions, and intermediate structures, such as middle *se* constructions, should show less lateralization, and less involvement of Broca's area, but more reliance on both hemispheres, as well as, possibly, more reliance on the subcortical structures of the brain. If syntax evolved gradually, through several stages, then it is plausible to expect that modern syntactic structures decompose into evolutionary primitives.

There are quite a few concrete and specific findings about how certain syntactic phenomena are processed. Consensus seems to be building around the claim that syntactic movement prompts additional activity in the left Inferior Frontal Gyrus (IFG) areas, clustering around the Broca's areas: BA 44 and 45, but also 46 and 47, according to some authors (see Ben-Shachar, Palti, and Grodzinsky 2004; Friederici *et al.* 2006; Grodzinsky 2010; Grodzinsky and Friederici 2006; Stromswold *et al.* 1996). In fact, the left IFG, especially Broca's area (BA 44/45), has been consistently linked to syntactic processing, including syntactic movement, across a wide range of approaches and methods (Shetreet, Friedmann, and Hadar 2009a,b; Friedmann 2006; Grodzinsky 2000, 2006; Friedmann and Shapiro 2003; Moro *et al.* 2001; Zurif 1995). There is converging evidence in the literature showing that increased syntactic complexity corresponds to increased neural activity in these specific areas of the brain (Caplan 2001; Indefrey *et al.* 2001; Just *et al.* 1996; Chesi and Moro, this volume, chapter 13). In assessing relative complexity, the literature on the topic typically uses as a starting point what are referred to as 'basic,' 'canonical' sentences, such as *John ran*, *Mary bought a book*. If the arguments in this chapter are on the right track, one would need to probe much below this 'basic' level, to the level of proto-syntax, in an attempt to compare the processing of such 'basic' finite sentences with the processing of proto-sentences.

Genetics is another area of great relevance to these considerations. A gene has recently been identified, the FOXP2, which is taken to play a role not only in articulation, but also in the processing of more complex (morpho-)syntax. The symptoms of KE family (who have a mutation) include simplified morphosyntax, implicating problems with functional categories and projections, including tense and TP. In an fMRI experiment (Liégeois *et al.* 2003), the unaffected KE family members showed a typical left-dominant distribution of activation involving Broca's area, whereas the affected members showed a more posterior and more extensively bilateral pattern of activation, as well as significant underactivation in Broca's area and its right homologue. According to Enard *et al.* (2002), there is evidence for positive selection of the gene by humans, which is clearly of relevance for evolutionary considerations.

However, at this point, linguists have no way of evaluating or making sense out of this great discovery in genetics, nor do we have any hypotheses to offer for any future endeavors in this field. If language evolved gradually, then in order to actively engage fields such as neuroscience and genetics, one needs to explore specific, theoretically informed, hypotheses about variability in linguistic complexity, as it relates to the stages of language evolution. Clearly, testing of such hypotheses is within reach.

6

Complexity in comparative syntax: the view from modern parametric theory

THERESA BIBERAUER, ANDERS HOLMBERG, IAN ROBERTS, AND MICHELLE SHEEHAN

6.1 Background

The theory of principles and parameters (P&P) as first proposed in detail in Chomsky (1981) and developed in versions of the Minimalist Program for linguistic theory offers a promising approach to the classical question of explanatory adequacy.* It is much less clear, however, whether this approach offers a way of approaching the question of complexity. This chapter aims to address this question: what (if anything) can modern parametric theory tell us about the formal complexity of grammatical systems?

In order to approach this question, we have to be clear about the nature of P&P. Its central idea can be summarized in (1), which essentially paraphrases Chomsky (1995):

(1) An I-language is an instantiation of the innate language faculty with options specified.

Here 'I-language' is taken in the sense of Chomsky (1986): the internal, individual language faculty characterized in intension by a generative grammar. The innate language faculty is that aspect of the human genome, apparently unique to humans, which makes the possession of an I-language possible, given appropriate environmental stimulus in early life. The theory of this faculty is Universal Grammar (UG). The 'options' of (1) are the parameters of UG, whose nature is the focus of much of the discussion below; suffice it to say for the moment that the parametric options

* This work is funded by The European Research Council Advanced Grant No. 269752: "Rethinking Comparative Syntax."

create the space of variation occupied, at least in part, by the typological diversity attested in actually occurring languages.

The P&P approach represents a major advance on earlier conceptions of language acquisition (see in particular the discussion in the Introduction to Chomsky 1981). Chomsky (1964) identified the goal of achieving explanatory adequacy in linguistic theory as accounting for the acquisition of a grammar in relation to Universal Grammar (UG). Earlier approaches, prior to 1981, had defined UG as a grammatical metatheory specifying a broad format for rules and some general conditions on rule application; a particular grammar as a system of language-specific, construction-specific rules; and language acquisition as rule induction, aided by an evaluation metric. This theory offered little hope for insights into either language typology or language acquisition. P&P stood in stark contrast to this from its inception. The leading idea was that UG contains an invariant set of principles associated with parameters which define the space of possible variation among the grammars of actual, individual I-languages. In these terms, language acquisition could be seen as setting the parameters of the native language on the combined basis of the innate UG and the triggering aspects of the primary linguistic data (PLD). In short, P&P appeared to significantly simplify the learning task, while at the same time providing typological insights in the form of 'parametric clusters'. Thus, it provided a way of connecting biolinguistics with language typology.

Despite its conceptual advantages over earlier approaches and its initial empirical promise in facilitating a new approach to typological questions, the P&P approach nonetheless has drawbacks, and these have gradually come to the fore in recent years.

The main difficulty encountered by P&P theory in recent years reflects an empirical issue. As research in comparative syntax has advanced, many more parameters than originally envisaged have been proposed to account for observed cross-linguistic variation. Much of this work has been quite successful, and there can be little doubt that our knowledge of the syntax of many of the world's languages and also of cross-linguistically recurring patterns has increased enormously since 1981. P&P theory has been an excellent heuristic. But Chomsky's criterion of explanatory adequacy requires more than this. Arguably, the direction that P&P theory has taken reflects the familiar tension between the exigencies of empirical description, which lead us to postulate ever more entities, and the need for explanation, which requires us to eliminate as many entities as possible. In other words, parametric descriptions as they have emerged in much recent work tend to sacrifice the explanatory power of parameters of Universal Grammar in order to achieve a high level of descriptive adequacy. The result is that the learning task remains mysterious, and the utility of P&P in solving this problem, which at the outset seemed so clear, is questionable.

Newmeyer (2004, 2005) was the first to construct a detailed critique of P&P theory, concluding that it was not living up to its promise. He advanced a number of criticisms of the approach, not all of which we agree with, and we do not endorse his conclusion that the approach should be abandoned. But Newmeyer (2005: 83) makes one extremely telling point: "... we are not yet at the point of being able to 'prove' that the child is not equipped with 7,846 [...] parameters, each of whose settings is fixed by some relevant triggering experience. I would put my money, however, on the fact that evolution has not endowed human beings in such an exuberant fashion." In other words, P&P theory places too much content in the innate endowment, and aside from general plausibility questions, this places an almost intolerable burden on any account of the evolution of language.

Finally, one of the most difficult problems for acquisition/learnability theory remains. This is often referred to as the *Linking Problem* (cf. Pinker 1984; Dresher 1999; and Fasanella and Fortuny 2013 for recent discussion). Parameters are defined over abstract linguistic entities, with the result that the language-acquiring child has to link these mental representations to actual physical entities in the speech signal. It is every bit as unclear in P&P theory as in almost any other approach to language acquisition how this happens.

Newmeyer's point holds in full force if learners must link and set large numbers of innately specified parameters, each independently of all the others. On these assumptions, the learnability problem takes its starkest form. Moreover, in this case, we might expect all grammatical systems to be equally complex. It would therefore appear on this view that P&P theory has little or nothing to say about the relative complexity of grammatical systems. If, however, parameters are interconnected in various ways, then this may simplify the learning task substantially. A concomitant of this is that the possibility then arises that certain parametric 'routes' to steady-state grammars are shorter—and hence in an obvious intuitive sense simpler—than others. This is the central idea that we try to develop in what follows.

Let us begin with a very simple example. It has been known since at least Kayne (1981) that languages vary as to whether they allow 'Exceptional Case-marking' infinitives or not. English allows this construction and French does not (in the canonical context involving *believe*-type verbs; see Kayne 1981):

(2) a. John believes Paul to write the best songs.
    b. *Jean croit Paul écrire les meilleures chansons.

We could therefore posit a parametric difference between English and French. It is also known that some languages have no infinitives at all, e.g. Modern Greek. In such languages, constructions corresponding to raising, control and ECM typically involve finite clauses in the indicative (complement to *believe*) or subjunctive (complement to *expect*) (D. Michelioudakis, p.c.).

So there are at least three options made available by UG: (i) both ECM and non-ECM infinitives (English); (ii) non-ECM, but not ECM infinitives (under *believe*-type verbs) (French); and (iii) no infinitives at all (Modern Greek). It is clear that the 'no infinitives' option obviates the need to choose between ECM and no ECM. In that straightforward sense, the third option is simpler than either of the other two: there is less for the learner to do. Moreover, we see that there are advantages from the learnability perspective in linking parametric options, and of course the more 'intrinsic' those links can be, as in this example, the better.

Moreover, the Minimalist Program (Chomsky 1995 *et seq.*) offers the possibility of seeing the nature of parameters in a new way, one which clearly offers a solution to the problem identified by Newmeyer. To see this, consider the three factors of language design put forward in Chomsky (2005):

(3) a. Factor 1: innate endowment (UG)
b. Factor 2: experience (PLD)
c. Factor 3: non-language-specific innate capacities

The first and second factors do not require much comment here and we note only that Factor 1, from a minimalist perspective probably contains far less than was assumed in former stages of the P&P approach. The 'third factors', according to Chomsky, include "(a) principles of data analysis that might be used in language acquisition and other domains; (b) principles of structural architecture and developmental constraints . . . including principles of efficient computation" (Chomsky 2005: 6). These factors clearly require further elucidation before the overall approach can be evaluated. Below, we will attempt to do this in relation to parametric variation and language acquisition. The general view that we take is that parametric variation is an emergent property of the interaction of the three factors listed in (3), and that parameters emerge as a consequence of the learning process. All that is pre-specified is (a) a small number of invariant properties of UG (first factor) and (b) general computational conservatism of the learning device (third factor).

As we hope to show, this view allows us to flesh out in interesting ways the question of whether languages differ in complexity. We address that question in §6.3 below. In §6.2, we set out in more detail our approach to parametric variation.

6.2 The nature of parameters: a proposal

6.2.1 *On the nature of parameters*

In this section, we introduce and illustrate the 'emergentist' approach to parameters just described. We must first state what does not vary, i.e. what is part of UG. UG determines the following properties of the linguistic computational system C_{HL}:

(4) a. certain formal features;
b. recursive, binary Merge;
c. a labelling algorithm;
d. Agree (feature-valuation, relating elements of syntactic structures).

Obviously much more needs to be said about all of (4a–d).[1] For present purposes, we take the class of formal features to include categorial features (±N, ±V, etc.), structural Case features (or equivalent), person, number and gender features (collectively φ-features), other features such as [±wh], [±neg], [±tense], etc, as well as purely diacritic features which simply trigger operations (different kinds of Merge, usually).[2]

Following Chomsky (1995: 243ff.), we take Merge to recursively combine two syntactic objects α and β to form a set {α, β}. The objects may be drawn from the Lexicon/Numeration (External Merge), or, if the members of an existing set {α, β} have internal structure, from within α or β (Internal Merge). The set formed by Merge requires a label K, i.e. Merge creates the object {K, {α, β}}. K is determined by either α or β, giving the effect of 'projection' of a syntactic category label, and hence endocentric structures. Finally, Agree involves valuing of formal features, which we take to be attribute-value pairs of the form [Person:3], [Number:plural], i.e. [Att(ribute):Val(ue)]. Features may enter the syntax without a value, i.e. as [Att:__], which the interpretative devices of the interfaces cannot read. Agree takes a pair of syntactic feature-bearing elements γ and δ such that for some feature F, one of γ and δ has the form [Att:__] and the other has the form [Att:Val]; the former is the Probe and the latter the Goal. The Probe must asymmetrically c-command the Goal and there must be no Goal' bearing an unvalued F such that the Probe asymmetrically c-commands Goal' and Goal' asymmetrically c-commands the Goal, i.e. the Goal must be the 'closest' possible Goal to the Probe. All of this is a fairly mainstream set of technical assumptions; for more details, see Chomsky (2001). The mechanisms described above are what we take to be the invariant core of UG.

Our principal departure from standard P&P thinking concerns the nature of parameters. Rather than taking them to be prespecified options of the kind 'A head X {precedes/follows} its complement YP', 'A head H, drawn from a set of heads L of licensing heads, formally licenses some element E in configuration C', etc., which are somehow genetically encoded, we take them to arise from **underspecification** of formal features in UG. This underspecification can take three forms, as follows:

[1] This exposition leaves two important issues open. First, the status of thematic roles, which, if some version of Baker (1988) is right, are structurally determined. It remains unclear whether the correlation between thematic role and relative syntactic position is determined by UG or emerges from some connection between event participation and structural prominence. Second, locality conditions, notably Relativized Minimality and the Phase Impenetrability Condition. We take Relativized Minimality to be a case of the general third-factor strategy of Minimal Search. Phase Impenetrability might be too. On the (im)possibility of connecting these two principles, see Rizzi (2013a).

[2] See Sigurðsson (2011) and Biberauer (2011, 2013) for the claim that all or many formal features may also be acquired on the basis of the interaction between a basic UG-given schema and the PLD. For present purposes, however, we keep to the rather more 'conservative' position which attributes some such content directly to UG.

(5) a. association of formal features with (functional) heads;
 b. values of formal features, triggering Agree;
 c. purely diacritic features triggering movement.

Certain heads are intrinsically potential bearers of formal features; this set may well be limited to the class of functional heads. For example, T bears φ-features of various kinds in many languages. In most Indo-European languages, T has Person and Number features and so we see agreement between the verb and the subject. Gender agreement between the verb and the subject is rare in Indo-European,[3] but found in many Semitic languages including Classical Arabic. Furthermore, as (5b) states, formal features may have their value specified or not; if they do not, then Agree is triggered, and it does seem to be the case that languages can vary as to the specific Agree operations they require (see Miyagawa 2010). The options in (5c) concern the distribution of the movement-triggering feature, which, following Biberauer, Holmberg and Roberts (2014, BHR henceforth) we write '^'. This sub-feature can be associated with any kind of syntactic dependency triggering Internal Merge, with languages differing in relation to which dependencies are associated with ^.

There is clearly a close relation between (4) and (5). In fact, (5) really says that a subset of the core properties of UG is optional in a given I-language; this is the content, for us, of Chomsky's statement in (1) that a given I-language is an instantiation of UG 'with options specified.' The 'theory of parameters' is nothing more than this: some subset of the universally available set of features is optional. In other words, to paraphrase Biberauer and Richards (2006), parametric variation emerges where UG 'doesn't mind.'

To elaborate slightly and give some more concrete examples of (5), (5a) includes such options as the mapping of features to heads, i.e. feature-scattering vs. feature-syncretism (see Giorgi and Pianesi 1997); presence vs. absence of features on heads; differing distribution and internal make-up of properties such as finiteness in clauses—this may underlie the variation between English, French and Modern Greek discussed above. Variation in Agree, and in its expression through inflectional morphology (which we take to be fairly closely associated to the presence of the features for learnability reasons; see Holmberg and Roberts 2013), gives rise to the differing properties of subject-agreement in English vs. Italian vs. Japanese, etc., and also to 'doubling' effects, e.g. Negative Concord, 'forked' modality in many South-East Asian languages (see Cheng and Sybesma 2003), 'bracketed' relative clauses (Peng 2011; Bradshaw 2009; Hendery 2012), etc. (5c) gives a range of movement options: V-movement in English vs. French (Pollock 1989) vs. Germanic verb-second languages (Holmberg and Platzack 1995); wh-movement in English vs. Chinese

[3] The Italo-Romance variety of Ripatransone is a rare case of this in Europe. See Ledgeway (2012: 299–310) for discussion and illustration.

(Huang 1982b); and, in combination with (5b), to different kinds of case systems including the distinction between ergative and accusative systems (Sheehan 2014); and variation in 'basic' head-complement order (BHR, Sheehan 2013a).

Essentially, (5) reduces to the statement in (6):[4]

(6) A given formal feature F may associate with a different set of heads (including the empty set) in different languages.

Here, for attribute-value features, 'F' ranges over [Att:val] and [Att:__]. To put things a little more formally, we can say that parameters involve generalized quantification over formal features, as follows:

(7) $Qh_{h \in P} [F(h)]$

Here Q is a quantifier, *h* is a head, P is the set of heads bearing the relevant formal properties (φ-features, movement-triggering features, etc.), and F is the set of formal features. Both F and P may be null in a given system, i.e. a given option may fail to apply.

This approach gives rise to the following informal taxonomy of parameters (Biberauer 2011; Biberauer and Roberts 2012a,b, 2013):

(8) For a given value v_i of a parametrically variant feature F:

a. **Macroparameters**: all (functional) heads share v_i;
b. **Mesoparameters**: all functional heads of a given naturally definable class, e.g. [+V], share v_i;
c. **Microparameters**: a small subclass of functional heads (e.g. modal auxiliaries, pronouns) shows v_i;
d. **Nanoparameters**: one or more individual lexical items is/are specified for v_i.

It is clear that the different kinds of parameters listed in (8) are hierarchically related to one another. So we are led to postulate different kinds of parameter hierarchies, whose relevance for acquisition we discuss below.

6.2.2 *Word order*

Roberts (2012) suggests the following parameter hierarchy for word order:

[4] See Lardiere (2008, 2009) for discussion of the language-specific assembly of features onto lexical items, and a proposal for the consequences of this view of cross-linguistic variation for the study of second language acquisition.

(9) ***Hierarchy 1: Word order***

a. Is head-final present?
- No: **head-initial**
- b. Yes: present on all heads?
  - Yes: **head-final**
  - c. No: present on [+V] heads?
    - Yes: **head-final in the clause only**
    - d. No: present on ...

Here we use the neutral term 'head-final'. More technically, head-finality may be instantiated as a complement-movement feature, following the general approach in Kayne (1994), or perhaps as a PF head parameter of the kind proposed by Richards (2004) and Sheehan (2013b); for present purposes we do not need to choose among these options. The higher nodes in this hierarchy define, first, rigidly head-initial systems and, next, rigidly head-final systems; in these systems all heads capable of varying in linear order in relation to their complements show a single, consistent order (we return below to the question of how a hierarchy structured as in (9) can be viewed as defining a learning path). These are macroparametric options both in the intuitive sense that they have massive effects in the grammars they determine, and in the sense defined in (8). The third output in (9) approximates to the typical Continental West Germanic situation (in which all clausal heads except C follow their complement); by the definition in (8), this represents a mesoparameter. Further 'down the hierarchy' on the unspecified lowest right branch, we define micro- and nanoparameters, ultimately specifying, for example, that in English the single lexical item *enough* follows rather than precedes the adjective it degree-modifies, unlike all other degree modifiers in English (i.e. *tall enough/*enough tall* vs. *very tall/*tall very*).

Roberts (2012) proposes that the parameter hierarchies arise from two interacting markedness conditions, Feature Economy (FE) (Roberts and Roussou 2003: 201) and Input Generalization (IG) (Roberts 2007: 273–275). These can be stated as follows:

(10) a. **Feature Economy** (FE):
Given two structural representations R and R' for a substring of input text S, R is less marked than R' iff R contains fewer formal features than R';

b. **Input Generalization** (IG):
If a functional head F sets parameter P_j to value v_i then there is a preference for similar functional heads to set P_j to value v_i.

Input Generalization plausibly follows from the acquirer's initial 'ignorance': not initially knowing what the categories in the target language are, the acquirer assumes an identified property/pattern to apply maximally generally; recognition that a new (sub)category needs to be distinguished, however, leads to re-evaluation of the initial input generalization. So-called *superset traps* are therefore circumvented because the child is assumed to be establishing the relevant inventory of syntactic categories incrementally (see Biberauer 2011, 2013; Branigan 2011, 2012 for more detailed discussion).[5] We take the conditions in (10) (perhaps along with the Subset Principle (Berwick 1985)) to arise from general cognitive optimization strategies applied to the task of language acquisition, *not* from UG. So the hierarchies are not part of UG, but determined by the underspecified parts of UG, interacting with conditions like those in (10) and the PLD. It is in this sense that parametric variation emerges from the three factors of language design given in (3). Since they do not form part of UG, the hierarchies cannot directly determine explanatory adequacy in Chomsky's (1964) sense. In fact, the hierarchies are descriptive taxonomies of the emergent system, i.e. epiphenomena. Since that system emerges from the interaction of the three factors in language design, and explicitly relates typological generalizations to language acquisition, and since the hierarchies aim to provide an explicit characterization of the way in which syntactic variation is structured, they obviously have explanatory value, though.

We can state things more precisely following the notation introduced in (7): given a head *h*, the set P of heads bearing the relevant formal properties (φ-features, movement-triggering features, etc.), and the set F of features, the general form of hierarchies and the steps in the learning path, as determined by FE and IG, will be as follows:

(11) a. Hypothesis I (ahead of any experience/analysis of PLD):
No head in P has F ($\forall h_{h \in P} \neg[F(h)]$); this hypothesis maximally satisfies FE and IG
b. Hypothesis II (at least one occurrence of F is detected in the PLD):
All heads in P have F ($\forall h_{h \in P}[F(h)]$); FE is overridden by PLD, IG is still satisfied
c. Hypothesis III (at least one non-occurrence of F is detected):
Some heads in P have F ($\neg\forall h_{h \in P}[F(h)]$); both FE and IG are overridden by PLD

The left branches of (9) reflect this ordering of (progressively weaker) hypotheses; (9/11a–b) reflect the **macroparametric** options; at the next level, generalization ranges over P'⊂ P, where P' is defined as a linguistically natural class (in (9), the class of

[5] This approach, then, can be characterized as falling into the class of maturational (rather than continuity) approaches to syntactic development (cf. Rizzi 1994 for discussion of the differences between these types of approach).

[+V] heads), and Hypotheses I-III are iterated over these classes; the shift from generalizing over P to generalizing over P' takes place since, at Hypothesis III, generalizing over P gives no clear outcome. FE and IG conspire to make each step refer to the minimal (FE) and the maximal (IG) proper subset of categories, hence the next level is the **mesoparametric** one. The **microparametric** level operates on still smaller subsets P"⊂P'⊂ P.The **nanoparametric** level operates on the smallest feasible subset (individual lexical items).

6.2.3 *Null arguments*

A further hierarchy, first put forward by Roberts and Holmberg (2010: 49), concerns null arguments. In terms of (11), we expect it to take the following form:

(12) a. Hypothesis I: no head in P (the set of probes) has uninterpretable φ-features.
b. Hypothesis II: all heads in P have uninterpretable φ-features.
c. Hypothesis III: some subset of P (the largest natural class P'⊂ P) has uninterpretable φ-features.

The system of hypotheses in (12) can be graphically illustrated by the diagram in (13), to which we have added a further mesoparametric option at the lowest level shown here:

(13) ***Hierarchy 2: Null arguments***

a. Are uφ-features present on probes?

No — **Radical pro-drop**

Yes — b. Are uφ-features present on all probes?

Yes — **Pronominal arguments**

No — c. Are uφ-features fully specified on some probes?

No — **Non-pro drop**

Yes — d. Are uφ-features fully specified on T?

Yes — **Consistent null subject**

No — …

Here 'radical pro-drop' refers to languages of the Chinese-Japanese type, which allow any pronominal argument to be 'dropped', and lack agreement inflections which could 'track' such arguments (see Huang 1984; Tomioka 2003; Saito 2007; Neeleman and Szendrői 2007 for discussion and differing analyses of this phenomenon). 'Pronominal argument' is intended in the sense put forward by Jelinek (1984): languages of this kind typically have very rich agreement marking for many, if not all, grammatical functions and a high degree of word-order freedom. Jelinek proposes that the agreement markers are the true arguments, incorporated into the verb from argument positions, with the optional realized nominal 'doubles' of these arguments being adjuncts, hence their somewhat free order (see also Speas 1990; Baker 1996). Again, these options are macroparametric, both in the clear sense that they have massively proliferating effects in the grammars they determine and in the sense defined in (8), with all heads sharing the same property. Question (c) corresponds to Hypothesis III in (11). At this point, then, mesoparametric options become relevant. This is arguably where the classical null-subject parameter comes in. This parameter refers to a sub-class of heads (finite T), as is usual for mesoparameters.[6] Also, we have changed 'present on probes' to 'fully specified on probes', since it is possible, for example, that [Person] is not present on T in English (its only putative instantiation is 3sg present *-s* which may be a default (Roberts to appear) or a [Number] morpheme (Kayne 1989)); if this is true, then English lacks fully specified T-probes. Why only T and not at least v becomes relevant at this point in this hierarchy is not yet clear to us, but it seems empirically correct that (a) subject agreement is cross-linguistically more common than object agreement and (b) (definite) null subjects are more common than null objects. Note, though, that the pattern of options is the same (features nowhere > features everywhere > features somewhere). The lower reaches of this hierarchy, not shown here, probably specify various kinds of partial null-subject systems, in the sense of Holmberg, Nayudu and Sheehan (2009), Holmberg (2010), in which φ-features are only partially specified on T.

6.2.4 *Word structure*

A further parameter hierarchy concerns word structure. We assume that complex words are formed by head-movement (cf. Baker 1988, 1996; Julien 2002). Hypothesis I then assumes no head-movement, i.e. highly analytic morphosyntax, while Hypothesis II assumes exceptionlessly instantiated head-movement, i.e. polysynthesis. Finally,

[6] It may seem odd to refer to finite T as a *class* of heads, but recall that it is necessarily present in all finite clauses and that it is also morphologically instantiated on all finite verbs and auxiliaries in a given language.

Hypothesis III assumes more limited head-movement, giving the mesoparametric options. This hierarchy is shown in (14):

(14) ***Hierarchy 3: Word structure***

a. Do some probes trigger head-movement?

- No: high analyticity
- b. Yes: Do all probes trigger head-movement?
  - Yes: generalized polysynthesis
  - c. No: do [+V] probes?
    - d. No: do all [+N] probes?
    - Yes: polysynthesis in the clause only

We take head-movement to be the formal operation which determines the internal structure of complex words. More specifically, following Roberts (2010), we see head-movement as a subcase of Agree, where the moved head has a proper subset of the features of the Probe (it is a 'defective Goal'). The notion of high analyticity referred to here is taken from Huang (2013), who sees exactly this property as a macroparameter characteristic of Chinese (of various kinds, including Modern Mandarin). Huang identifies a range of syntactic effects of this property, alongside extreme morphological analyticity, which are construed as showing lack of head-movement in the system as a whole. The next option is polysynthesis, as discussed and analyzed by Baker (1996). Following Baker, we take 'the polysynthesis parameter' to involve head-movement as a highly prevalent property of the system, affecting all lexical categories. Both analyticity and polysynthesis are again macroparameters: they have massive effects on the systems they determine. Fully polysynthetic languages, i.e. those which show polysynthesis both in the clausal and the nominal domains fall under Hypothesis II at the macro-level. A positive value for [+V] at the mesoparametric level would give rise to a language which is polysynthetic only in the clausal domain (e.g. Michif; Bakker 1997). Lower in the hierarchy we encounter parameters determining familiar cases of head-movement; Biberauer and Roberts (2012a) propose a sub-hierarchy for verb-movement which encompasses a meso-parametric option for V-to-T movement (as in Pollock 1989), microparametric options for auxiliary-movement (encompassing Modern English '*have/be*-raising' (Emonds 1976)), and nanoparameters affecting individual auxiliaries in certain varieties of Modern English. N-movement of the kind discussed by Bernstein (1991) and Longobardi (1994) defines a further mesoparametric sub-hierarchy.

True macroparameters sit close to the top of each hierarchy, as here all heads parametrized for the feature(s) in question behave as one. Moving down the hierarchy, parameters become more 'micro', behaving in a non-uniform, differentiated fashion. Crucially for present purposes, microparameters are also inherently more complex than the systems defined higher in the tree. This is simply because a smaller number of parameter settings are needed to give rise to the higher options, as we have seen, while systems defined lower down in the hierarchy require more parameters and apply first to natural subsets of the entire set (mesoparameters), to smaller classes of functional categories F (microparameters), and ultimately to single lexical items (nanoparameters). Moving 'down a hierarchy', then, systems become more marked, having a longer and more complex description than the higher options. The possibility also arises that lower options are further along a given learning path. Here it is worth noting that many, perhaps all, nanoparametric options fall outside the core system defined by the hierarchies under discussion here. To the extent that nanoparametric options involve high-frequency elements, they appear to be acquired as independent lexical items, independently of the more general properties of the system to which they belong; hence the much-discussed U-shaped acquisition pattern associated with the acquisition of high-frequency irregulars (cf. Marcus *et al.* 1992 for detailed discussion). In our terms, forms of this type would therefore not be acquired as a result of progressing down a given hierarchy, although their connection to specific hierarchies—in the sense that they appear to represent isolated instantiations within a given system of a pattern that can be seen to hold more systematically in other systems—is clear. We return to this point in section 6.3 below.

6.2.5 *A'-movement*

The fourth hierarchy is more tentative than the others, and we introduce it largely to illustrate how our approach can shed light on the general question of grammatical complexity. It concerns what, following É.Kiss (1995), we can loosely refer to as 'discourse configurationality'. More technically, it concerns options of A'-movement. As such, two ingredients are crucial: the concept of phase, as introduced in Chomsky (2000) and developed in Chomsky (2001), and A'-related formal features (for simplicity, we will refer to [focus], [wh] and [topic]). We assume, in part following Chomsky (2007), that phase-heads define local domains, license movement to and/or through their left periphery, and trigger A'-movement. Suppose that C, D and v are phase-heads (there may of course be others). Suppose further that there is universal functional pressure for systems to encode focalization/topicalization, these being a component of the 'second' type of semantics Chomsky highlights in referring to 'duality of semantics' (we return to this point below). Formally, let us assume that elements which are to undergo focalization/topicalization and A'-movement more generally will be 'inflected' to reflect this fact, i.e. they will differ from elements which can remain in situ in virtue of bearing one or more A'-features of the relevant kind. Cross-linguistic investigation has shown that A'-movement is often to left-peripheral positions within CP, vP and DP. At the same time, syntactic locality (subjacency/ island conditions) severely restricts movement to the left periphery, forcing all (long-distance) movement to be successive-cyclic. Phase-heads can function as escape hatches (licensing cyclic movement *through* their left periphery, without interpretive effect) or as targets (licensing movement *to* their left periphery, giving an appropriate discourse interpretation). Under certain circumstances, this phasal 'escape hatch' is not available, however. Let us suppose that all phase-heads can, in principle, allow successive-cyclic movement *to* their edge and, where they do not represent the last-merged phase-head in the clausal domain, also *through* their edge (this is, then, an extension of Chomsky's 1973 proposals regarding the successive-cyclicity of *wh*-movement). The *through* options available to non-last-merged phase-heads are clearly restricted where island effects are observed. Two considerations which appear to be relevant in determining the possibility of escaping from phasal domains are (i) the relevant domain having been spelled out (which we take to mean that its internal structure has become invisible to the computational system, with the result that it cannot be targeted by either of the operations Agree or Move) and (ii) the relevant domain having been 'sealed off' by a highly specified nominal head whose rich featural specification precludes the possibility of other elements being extracted across it (i.e. Relativized Minimality considerations of the type discussed *i.a.* in Starke (2001) and Rizzi (2001, 2013)). Precisely which *to* and *through* options are permitted and whether a given system includes a nominal head of the relevant kind we assumed to be a matter of parametric specification (the distinction between preposition-stranding and obligatory pied-piping would presumably be determined by this kind of option applying to PP). More specifically, consider (15):

(15) ***Hierarchy 4: A'-movement***

a. Do phase-heads trigger A'-movement?

No: UNATTESTED

b. Yes: Can all phase-heads trigger A'-movement *to* their edge?

Yes: **'free word order' (a)**

c. No: Is C unable to trigger *wh*-movement *to* its edge?[7]

Yes: ***wh*-in-situ+scrambling (b)**

d. No: Does v restrict *wh*-movement *through* its edge?

Yes: **syntactically ergative languages (c)**

d. No: Does v restrict [focus]- and [wh] movement *to* its edge?

Yes: ***wh*-movement + scrambling (d)**

e. No: Does v restrict [focus] and [topic] movement?

Yes: ***wh*-movement only(e)**

No: ...

Here, we see that one of the options given by the broadest question, namely that of foregoing A'-movement, is in fact a non-choice. We return to the matter of such no-choice parameters below. Type (a) languages include Warlbiri and many other Australian languages, Latin, the Slavonic languages and others. These languages have very liberal scrambling, both to the *Mittelfeld* and to the left-periphery, and also subextraction from nominals, creating the possibility that adjectival and other adnominal modifiers can appear somewhat distant from the noun they modify, one characteristic often thought to characterize 'free word order'. This type of language we assume to be the reflex of a formal system in which all phase-heads (and relevant

[7] This negative formulation (15c) aims to highlight the fact that C only attracts a subset of A'-elements to its edge, [focus]- and [topic]-bearing ones. Given the prominence of *wh*-interrogatives in the input, it is plausible to assume that acquirers will register *wh*-in-situ structures, with the result that they will conclude that a specific phase head, C, does not trigger movement of *wh*-elements and that, therefore, A'-movement, which had initially been assumed to be possible across-the-board, is (potentially) restricted to topics and foci. The proposal is therefore *not* that the child is 'interrogating' the input on the basis of the c.-question.

clause-internal elements; see below) have the possibility of being specified for one or more of [topic], [wh] and [focus] with one or more associated ^-features, triggering movement, and in which the highly specified island-creating nominal head mentioned above is absent. More specifically, C, v and D in languages of this type will all be able to trigger both movement *through* their edge (by virtue of their being able to bear an independent ^, not associated with a specific substantive formal feature) and topic-, *wh*- and focus-movement *to* their edge (by virtue of their ability to bear [topic], [wh] and [focus] features respectively alongside ^). Thus systems of this kind can be thought of as instantiating a macro-option in relation to phase-heads in that they treat all the moving *to* and *through* options associated with these elements identically. This means that fewer types of C, v and D (i.e. fewer sub-categories) need to be acquired, as one would expect for an option located high on a hierarchy defining a learning path.

Type (b) languages include Japanese and Korean; these languages have quite liberal scrambling, but no clausal-level overt *wh*-movement in interrogatives. These phenomena we view as indicative of the fact that C, v and D cannot be treated identically, as was the case of Type (a) languages. Instead, it seems to be necessary to distinguish between the way in which C, v and D bear [topic]- and [focus]-features on the one hand and [wh]-features on the other. Consequently, languages of this type require the postulation of a larger number of distinctly specified heads than Type (a) languages.

Type (c) languages include Tagalog and many other Polynesian languages. Strikingly, these are ergative languages, which restrict *wh*-extraction to absolutive-marked arguments. In terms of analyses like Aldridge (2004), Coon, Mateo Pedro and Preminger (2011) and Sheehan (2014), this restriction entails that only arguments that can be targeted for movement *through* the edge of vP by virtue of the fact that they are not first-merged within that edge can in fact be extracted.[8] The properties of Type (c) are clearly also in part determined by aspects of Hierarchy 5, which we discuss below, pointing to the fact that the hierarchies may interact with one another (a topic which we must leave aside here).

Type (d) includes German and Dutch, i.e. systems which feature *Mittelfeld* scrambling, and overt *wh*-movement. In formal terms, these are systems in which C and (non-island-inducing) D may be specified as for Type (a) and (b) languages, but where v cannot bear ^-associated [focus] or [wh]; thus only [topic]-elements can remain within the vP-edge (scrambling), while [focus] and *wh*-elements may move *through* this edge to CP.

[8] These analyses propose that movement of the absolutive to the outer specifier of vP serves to trap the transitive (ergative) subject inside that phase. As Assmann *et al.* (2012) show, this restriction affects only transitive ergative subjects and not other arguments inside vP, suggesting that what blocks extraction of the transitive subject is its base-generation in the phase edge.

Type (e) includes English, North Germanic and the Romance languages, which permit little or no scrambling, but do feature overt *wh*-movement.[9] In these languages, C and (non-island-inducing) D are specified as for Type (a), (b) and (c) languages, but v is not associated with substantive features of any kind; it only bears ^, serving as an escape hatch for movement on to C.

6.2.6 *Alignment*

The fifth and final hierarchy concerns alignment, in the general sense of how the core grammatical functions are marked in the case/agreement system. Here we present a version of this hierarchy proposed in Sheehan (2014):

(16) ***Hierarchy 5: Case and alignment***

a. Does transitive v assign θ-related case (ERG) to its specifier in L?

- No: Accusative
- b. Yes: Do all v in L assign ERG?
  - Yes: Split S (Chol, Basque)
  - c. No: Does v_{ERG} bear an EPP feature in L?
    - No (morpho ergative: Warlbiri)
    - d. Yes: Does v_{ERG} assign structural Case in L?
      - Yes: Low ABS (Tagalog)
      - No: High ABS (Dyirbal)

The first option distinguishes the familiar accusative alignment, found covertly in English, overtly in Latin, Russian, Japanese, etc., from all non-accusative systems. The second parameter separates split-S languages, also known as *stative-active languages*, which show ergative alignment only with the single argument of an unaccusative verb (cf. Mithun 1991; Laka 2006). The third distinguishes languages in which ergative alignment is purely a matter of case and/or agreement marking (cf. Anderson 1976) from those which disallow the A'-extraction of ergative-marked DPs (a property which has come to be known as 'syntactic ergativity', cf. Coon *et al.* 2011). The final parameter concerns the source of absolutive case and hence the extent to which the absolutive argument shows 'subject properties' of various kinds (ability to be controlled in non-finite clauses, absence in non-finite contexts). In transitive clauses, then, the internal argument can display these properties in High-ABS languages, because Absolutive is uniformly assigned by T (cf. Legate 2008, 2011).

9 Though Spanish may have scrambling in marked VOS orders (see Ordóñez 2000; Gallego 2013).

Hierarchy 5 departs a little from the form of the earlier hierarchies. The hierarchies in (12–15) all have at the highest node the question of whether the relevant property is instantiated in the system at all (the 'head-final' feature, φ-features, head-movement and A'-movement respectively). In this way, the highest option maximally satisfies both EF (no feature) and IG (generalization of the absence of the feature). It would be possible, obviously, to add a macro-parametric option to the top of hierarchy 5, determining whether structural Case—and therefore A-movement—is present in a given language (see Diercks 2012). As the parameterization of structural Case remains somewhat controversial, though, we leave this option open here. This minor difference aside, the five parameter hierarchies provide a fairly rich characterization of the grammar of natural languages and, we argue, create new possibilities regarding the calculation of grammatical complexity.

6.3 Complexity and emergent parameters

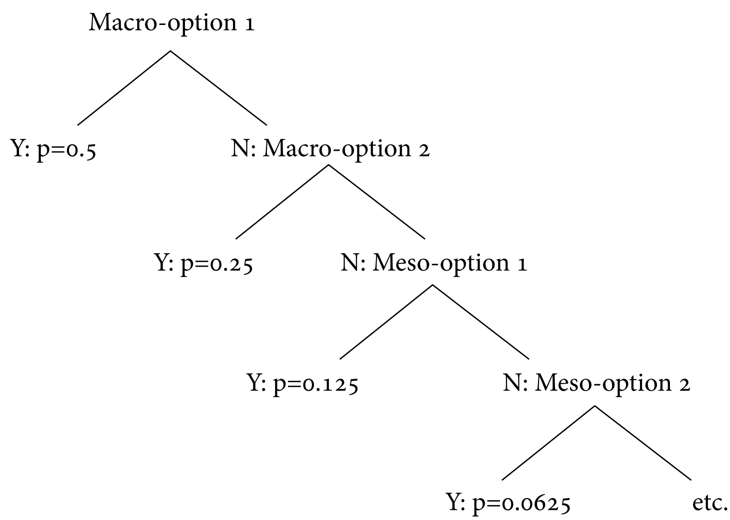
Consider now how the parameter hierarchies presented above might contribute to an understanding of grammatical complexity. First, the lower positions in the hierarchies correspond to more micro-parametric options: going down a given hierarchy, we move from macro- to meso- to micro-variation (as noted above, nanovariation is lexically idiosyncratic and thus in a sense outside the hierarchies). Second, the lower options behave in a non-uniform, differentiated fashion which is inherently more complex than the systems defined higher up. Third, each parameter hierarchy can be thought to define a learning path, much in the sense of Dresher (1999), with the higher options inherently preferred by the acquirer, because Input Generalization favors the higher options in the absence of PLD regarding more specified options. Finally, where hierarchically lower options rely on low-frequency components of the input, we predict Input Generalization to lead to overgeneralization, which may, in turn, lead to the loss of such options, resulting in a less complex system. Essentially, highly irregular 'low' options will either be lost or 'analogized out of the system' over time (note in this connection the close similarity between Input Generalization and the neo-grammarian notion of analogy).

There are at least two distinct ways to calculate the notion of complexity using hierarchies and we will now consider these in turn.

6.3.1 *Complexity and probability*

One way of calculating parametric complexity involves equating complexity with probability. We reason as follows: all else being equal, there should be a roughly 50/50 chance of taking a given option at each independent choice point, making lower positions in the hierarchy cumulatively less probable. We can quantify the probability associated with a given output of the hierarchy as 0.5^n, where n is the level of embedding in that hierarchy. (17) provides a dummy hierarchy with dummy relative probabilities:

(17)



We look at a language's position in each of the five hierarchies (to the extent that this is feasible), assign a value for $p = 0.5^n$ in each case, and then give the product of the five independent probabilities. This gives a complexity index for (the syntax of) each language, equivalent to the probability of this grammar. The smaller this value, the more complex the grammar of the language.

Several interesting points immediately arise concerning this way of reasoning, in conjunction with the fact that we have posited five interacting, but independent hierarchies. The first point is that it seems unlikely that any system will be maximally unmarked (in featural terms). To be maximally unmarked would entail being, as it were, 'at the top' of all five hierarchies. Although this would be the preferred option in terms of maximal satisfaction of both EF and IG, it may be ruled out for independent reasons. Consider what the properties of the least-marked possible system would have to be:

(18) a. Harmonically head-initial;
b. Radical pro-drop;
c. High analyticity;
d. No A'-movement (i.e. no mechanism of focusing, topicalization, *wh*-movement, scrambling)
e. Accusative alignment (or no Case if hierarchy 5 is expanded upwards)

We conjecture that no language has a system of this kind. Thai, Vietnamese and possibly other South-East Asian languages come close, but all of them, to our knowledge, show some word-order disharmony (final modals, some head-final orders within the nominal; see *i.a.* Duffield 2001; Enfield 2003; and Simpson 2005) and they also permit information structure-related A-bar movement (cf. Hinds 1989; Phimsawat 2011; Badan and Del Gobbo 2011). Many creoles also come close, but all feature *wh*-movement. If the maximally unmarked system were found, the prediction is that it would presumably represent a 'basin of attraction' in that it would be impossible—or at least extremely difficult—for such a system to change, for the reasons we discussed above in relation to the diachronic conservativity of macro-parametric settings. An important question, then, is why such systems do not seem to exist.

At this point, functional considerations come into the picture. We propose that certain options, which are left open by UG in principle, are impossible in practice for functional reasons (cf. Biberauer, Holmberg, Roberts, and Sheehan 2010; Biberauer, Roberts, and Sheehan 2014; Sheehan 2014; and Biberauer 2011, 2013). This is perhaps clearest in the case of the least-marked options in Hierarchy 4. As we saw, to be consistent with our general markedness conditions and with the first three hierarchies, the most unmarked system here, as dictated by FE and IG (see (10)), is that in which there are no A'-movement triggers at all. We take it that UG in principle allows such an option, but that functional considerations rule it out of the parametric 'gene pool': no system entirely lacking a formal means to focalize/topicalize constituents is likely to survive as it falls short of basic expressivity needs—it arguably undermines one of the two types of semantics in Chomsky's 'duality' (cf. Fortuny 2010; Biberauer 2011). Since UG, as a formal system, is entirely indifferent to questions of expressivity, the formal options exist, but, for reasons to do with thought, conceptualization and communication—i.e. the cognitive systems UG interfaces with—they are never instantiated. Parameters which offer this non-choice are referred to as 'no-choice' parameters. Another important factor is contingent pressure on languages from language contact. While we take contact-induced change to be constrained by UG and its interaction with more general non-language-specific considerations, such change may follow a different path from internally-triggered change. As such, marked systems can come into existence out of less-marked systems because of (i) functional pressures and/or (ii) language contact.

We are now in a position to address the central question of this chapter: do grammars differ in complexity and, specifically, do they differ in terms of complexity defined in terms of parametric probability? To answer this, we propose a thought experiment (in advance of the real experiments, which are the object of ongoing work). Applying the formula discussed above, as we go down a given hierarchy, the probability of being assigned a given parameter value decreases as a function of depth: $p = 0.5^n$ (where n = level of embedding; cf. (18)). We can then calculate the probability of a given language by multiplying together all these independent probabilities (assuming, for illustrative purposes, that these probabilities are indeed independent).

So let us see how this works for the grammars of some fairly well-known and well-studied languages across a reasonable typological, genetic and areal range. We leave nanoparameters aside, since it is clear that the degree of complexity added by elements with idiosyncratic formal specifications does not seem amenable to the kind of regular, level-of-embedding-based quantification we are proposing for macro-, meso- and microparametric properties. In this respect, nanoparameters may be just 'noise' from the point of view of computing overall complexity. Moreover, considerations such as frequency need to be taken into account in some way when calculating the complexity added by irregulars of different types. We leave these questions aside for the present, focusing on parametric options that seem more amenable to quantification on the basis of the hierarchies that we have been discussing.

First, ***English*** is: (basically) harmonically head-initial (0.5 on Hierarchy 1), non-pro-drop (0.125 on Hierarchy 2), shows Aux but not V-movement (0.03125 on Hierarchy 3),[10] has *wh*-movement but no scrambling (0.03125 on Hierarchy 4), and is accusative (0.5 on Hierarchy 5). The product of these probabilities is 0.003%, making English a relatively complex language.

Consider next ***Mohawk*** (here our information comes from Baker 1996 and the references given there). Leaving aside head-initiality/finality for a moment for a reason that will immediately become clear, this language has pronominal arguments (0.25), polysynthesis (0.25), free word order (0.25), and split-S alignment (0.25). Baker (1996) argues extensively that it is impossible to ascertain the nature of head-complement order owing to the language's pronominal-argument, polysynthetic nature, which has the consequence that all nominal arguments, both in the clause and inside the nominal (e.g. possessors) are adjuncts which can appear either left- or right-adjoined to the clause/nominal. If this is true,[11] then we can infer that the basic word-order parameter is never set. This scenario is arguably problematic in the

[10] Biberauer and Roberts (2012a) tentatively place the Modern English Aux-movement option at the 3rd level of embedding in the verb-movement hierarchy they propose. But this hierarchy must be further embedded in Hierarchy 3. If it is embedded at the deepest right branch in (14), then the English option would be at the 5th level of embedding. For expository purposes, that is what we assume here.

[11] One might question this conclusion on the basis of the fact that Mohawk features complement clauses that seem rather similar to English *that*-clauses, both in respect of the fact that they systematically surface postverbally and in respect of their being introduced by an optional complementizer-like element, *tsi* (cf. Ikeda 1991 for further discussion).

context of traditional parametric approaches; in the context of the emergentist approach argued for here, however, it simply entails that no question ever arises regarding the presence of head-finality in the system, with the consequence that the word-order parameter is set to head-initial, giving a complexity value of 0.5 in this domain. The product of these probabilities is 0.195%. Hence we see that Mohawk is, perhaps surprisingly, somewhat less complex than English.

Mandarin Chinese (Huang 1982b, 2013; Huang, Li and Li 2009) is harmonically head-final in [+N] but not in [+V] (0.0625), radical pro-drop (0.5), highly analytic (0.5), has topicalisation to the left-periphery (Badan and Del Gobbo 2011), scrambling (Soh 1998), no *wh*-movement (0.125), and accusative alignment (0.5), so the overall probability for this language is 0.098%, meaning that Mandarin falls somewhere between English and Mohawk in terms of complexity (calculated in terms of probability).[12]

Japanese is harmonically head-final (0.25), radical pro-drop (0.5), agglutinating in both verbal and nominal domains (0.5; see Julien 2002 and Neeleman and Szendrői 2007), *wh*-in-situ + scrambling (0.125) and accusative (0.5). This gives an overall complexity index (probability) of 0.391%, making Japanese less grammatically complex than Mohawk.

Finally, ***Basque*** is harmonically head-final (0.25), has pronominal arguments (0.25), is agglutinating (0.5), has *wh*-movement+scrambling (0.125) and split-S alignment (0.25), giving a complexity index of 0.098%, identical to that of Mandarin Chinese.

6.3.2 *Complexity indices*

A further method for calculating the grammatical complexity of a language using the hierarchies involves assigning each output a complexity index directly, based on the number of choices it entails and taking the average across all five hierarchies. This approach distinguishes the notion of complexity from that of probability, and interestingly gives a slightly different picture for the languages under discussion. The following diagram indicates the complexities associated with the various outputs of a binary branching hierarchy of the kind we have been discussing (the higher the index the higher the complexity):

[12] There is a question whether a language which lacks φ-features can be considered to be accusative. It is possible that Hierarchy 5 is not even activated in a system which lacks φ-features and Case. We leave this matter to one side here.

(19)

- Macro-option 1
  - Y: c=1
  - N: Macro-option 2
    - Y: c=2
    - N: Meso-option 1
      - Y: c=3
      - N: Meso-option 2
        - Y: c=4
        - N : c=4

We can now use these complexity indices to calculate the average grammatical complexity of a given language, where this time, the *higher* the number, the more complex the language. Consider first ***English***. It is (basically) harmonically head-initial (c=1 on Hierarchy 1[13]), non-pro-drop (c=3 on Hierarchy 2), shows Aux, but not V-movement (c=5 on Hierarchy 3), has *wh*-movement but no scrambling (c=5 on Hierarchy 4), and is accusative (c=1 on Hierarchy 5). This gives an average complexity index of 3 across the five hierarchies.

Now consider **Mohawk**, which, as discussed above has pronominal arguments (c=2), polysynthesis (c=2), free word order (c=2), split-S alignment (c=2), and head-initial word order (c=1). Mohawk thus has an average complexity index of 1.8, again slightly less complex than English.

Applying the same methodology to Mandarin, Japanese and Basque gives the following overall picture:

(20) Japanese: 1.6
Mohawk: 1.8
Mandarin: 2
Basque: 2
English: 3

[13] If the highest option in a hierarchy does not in fact involve an explicit choice, as suggested above, it may not be correct to assign this option a choice-count of 1. We leave this matter aside here.

Compare the results from the probability-based calculations:

(21) Japanese: 0.391%
Mohawk: 0.195%
Mandarin: 0.098%
Basque: 0.098%
English: 0.003%

We observe, then, that, while the overall complexity indices are different, the relative complexities are identical whether we calculate complexity by probability or choice-count. We stress that these calculations are intended as purely illustrative, as a 'proof-of-concept' of the idea that our parameter hierarchies can yield ways of quantifying the complexity of a given grammar. At this stage, we are comparing just the relatively 'large-grained' properties that are the focus of our current research.

A couple of comments are worth making here. First, we can observe that there are no extreme outliers (except possibly English if probabilities are used). This is an encouraging outcome, although of course this conclusion is extremely tentative, being based on just five languages. Second, English emerges, in both instances, as the most complex language, which is not surprising, in particular given the contribution made by what is known to be a very unusual and marked property of this language, namely its auxiliary system.[14] Third, Japanese seems remarkably simple if we compare it to English, Basque and Mandarin in particular. In connection with Japanese, it is worth pointing out that harmonically head-final word order, radical pro-drop, agglutinating morphology, *wh*-in-situ with scrambling and accusative alignment are properties of a large number of languages spoken across the northern part of Asia (essentially the allegedly 'Altaic' languages, comprising Japanese, Korean, and the Turkic, Mongolian and Tungusic families; Menges 1975). We do not wish to assert that this supports any version of the Altaic hypothesis, but it is possible that the regional prominence of these properties may reflect the fact that they are relatively unmarked properties which are therefore of the kind that we might expect to spread readily through contact. Nichols (1992) identifies North Asia as one of the world's major 'spread zones', i.e. areas in which the topography permits extensive and innovative language contact (see also Dryer 1998).

A final point is that of course here we are only looking at (morpho-)syntactic complexity; it would be revealing to carry out a similar exercise in relation to phonological properties, whereafter the further questions naturally arise of whether and, potentially, how phonological and morphosyntactic complexity might be

[14] Of course, the measured complexity of English could also be an effect of the fact that it is the best-studied language in generative grammar. Note crucially, though, that an English bias might lead us to expect it to be the *least* marked system, contrary to fact. In this way, the result is doubly encouraging.

combined to give an overall picture of the formal complexity of language systems. But that would go well beyond the scope of this chapter.

6.4 Conclusion

We have tried to illustrate in the foregoing a new approach to parametric variation, which, we believe, has a number of virtues. First, it overcomes the earlier objections of Newmeyer (2005) to classical parametric theory. Second, it can be reconciled with certain types of functionalist approaches (cf. the notion of 'no-choice' parameter introduced in §6.2). Third, it makes new diachronic, typological and acquisitional predictions and fourth, it is fully compatible with minimalist assumptions on language design, arguably allowing certain aspects of Chomsky's third factor to be made more precise. Finally, as we saw in §6.3, our approach may allow for an overall quantification of the complexity, and hence perhaps the markedness, of grammars. For all of these reasons, we believe our approach to be worth considering and developing further.

We close with a final remark on terminology. It is possible that the term 'parameter' may no longer really be appropriate, as the sense in which we understand it is fairly different to its original sense in Chomsky (1981). However, two considerations lead us to retain the term. First, introducing a new term would almost certainly create unwelcome terminological confusion. Second, we see this work as maintaining the spirit of the principles and parameters approach. It should not be forgotten that technical scientific terms change their denotations, both their extension and their intension, as knowledge progresses. In this respect, the ways in which the denotation of 'syntactic parameter' has changed since 1981 should be seen as a sign of progress.

7

The complexity of narrow syntax: Minimalism, representational economy, and simplest Merge

ANDREAS TROTZKE AND JAN-WOUTER ZWART

7.1 Introduction

The issue of linguistic complexity has recently received much attention by linguists working within typological-functional frameworks (e.g. Miestamo, Sinnemäki, and Karlsson 2008; Sampson, Gil, and Trudgill 2009). In formal linguistics, the most prominent measure of linguistic complexity is the Chomsky hierarchy of formal languages (Chomsky 1956), including the distinction between a finite-state grammar (FSG) and more complicated types of phrase-structure grammar (PSG). This distinction has played a crucial role in the recent biolinguistic literature on recursive complexity (Sauerland and Trotzke 2011). In this chapter, we consider the question of formal complexity measurement within linguistic minimalism (cf. also Biberauer *et al.*, this volume, chapter 6; Progovac, this volume, chapter 5) and argue that our minimalist approach to complexity of derivations and representations shows similarities with that of alternative theoretical perspectives represented in this volume (Culicover, this volume, chapter 8; Jackendoff and Wittenberg, this volume, chapter 4). In particular, we agree that information structure properties should not be encoded in narrow syntax as features triggering movement, suggesting that the relevant information is established at the interfaces. Also, we argue for a minimalist model of grammar in which complexity arises out of the cyclic interaction of subderivations, a model we take to be compatible with Construction Grammar approaches. We claim that this model allows one to revisit the question of the formal complexity of a generative grammar, rephrasing it such that a different answer to the question of formal complexity is forthcoming depending on whether we consider the grammar as a whole, or just narrow syntax. The grammar as a whole, including interface components in addition to narrow syntax, as well as recursive interaction among subderivations, is vastly more complicated than a finite-state grammar, but there is no reason for concluding that narrow syntax is not simply finite-state.

The chapter is structured as follows. In section 7.2, we clarify the generative perspective on measuring linguistic complexity and distinguish this approach from performance-oriented notions that are traditionally confused with the generative account. Section 7.3 shows how complexity reduction in the domain of syntactic representations results in points of convergence between minimalism and other perspectives that assume that linguistic complexity does not arise from syntax alone. In section 7.4, we turn to the layered-derivation perspective on linguistic complexity and argue that narrow syntax can be captured by a finite-state device and, therefore, falls low on the Chomsky hierarchy. Section 7.5 summarizes the main results of the chapter, followed by a short conclusion.*

7.2 The generative perspective on measuring linguistic complexity

Within the generative paradigm, the question of the comparative complexity of languages, as discussed in recent typological-functional literature, does not arise (cf. also Sauerland 2014). It has been a core assumption of the generative research program since its beginnings that the complexity of individual languages is determined by the invariant biological mechanisms underlying human language in general. Of course, a more differentiated picture might emerge when we look at the interconnection of specific parameters that are set in different languages and assume that there is some complexity metric that classifies parametric 'routes' to certain grammars as simpler than routes to certain other grammars (Biberauer *et al.* this volume, chapter 6). However, within generative linguistics, such an approach is not uncontroversial, since the theory of grammar, as currently understood, has no room for connecting statistical generalizations to properties of the faculty of language in the narrow sense ('I-language') or to any principles and parameters associated with I-language.[1] As Newmeyer (2007: 240) points out, "correlations between complexity and rarity are not [. . .] to be expected, since implicational and frequency-based generalizations do not belong to the realm of I-language." In this chapter, we do not

* We thank the audiences at the workshop on *Formal Linguistics and the Measurement of Grammatical Complexity* (23–24 March 2012, University of Washington, Seattle) and at the workshop on *Complex Sentences, Types of Embedding, and Recursivity* (5–6 March 2012, University of Konstanz). We are especially grateful to the editors of this volume, Fritz Newmeyer and Laurel Preston, for the careful discussion of every aspect of this chapter. All remaining errors and shortcomings are our own. Andreas Trotzke gratefully acknowledges financial support from the DFG Excellence Initiative (University of Konstanz, project no. 65411).

[1] In this chapter, we equate I-language with the faculty of language in the narrow sense as defined in Hauser, Chomsky, and Fitch (2002), i.e. the component of the grammar that applies simple rules merging elements, also referred to as 'narrow syntax' below.

delve into these issues. Instead, we focus on basic derivational and representational aspects of linguistic complexity that are not subject to variation.[2]

If we concentrate on I-language, the issue of measuring complexity among natural languages disappears. What does not disappear, however, is the question where the grammar of natural language falls on the complexity hierarchy of *formal* languages. In the 1950s, Chomsky showed that a particular type of recursion is essential to drawing the line between the phrase structure models of language that he proposed and models of language prevalent in contemporary structuralist thinking. In particular, Chomsky (1956, 1959) showed that self-embedding involves the kind of complexity that requires (at least) context-free grammars, rather than less complex types of grammar (specifically, finite-state devices). Chomsky (1959: 148) defined this notion of self-embedding as follows (*I* is the identity element, i.e. zero, and $\Rightarrow$ indicates a derivation involving rewrite operations):

(1) A language L is *self-embedding* if it contains an A such that for some φ, ψ ($\varphi \neq \mathrm{I} \neq \psi$), $\mathrm{A} \Rightarrow \varphi\mathrm{A}\psi$.

The definition characterizes as self-embedding any language that contains a string A and allows the derivation from A of a string that properly contains A, that is, A is preceded and followed by two non-trivial strings. Chomsky (1957) went on to show that patterns such as (2) exist in English (slightly modified from Chomsky 1957: 22, with *It's raining* as the declarative sentence S). These patterns clearly satisfy the definition of self-embedding in (1):

(2) a. S $\Rightarrow$ If S, then it's true.
b. S $\Rightarrow$ If it's raining, then it's true.
c. S $\Rightarrow$ If if it's raining, then it's true, then it's true.
d. (. . .)

As Chomsky notes, S in (2a) can in fact have the same structure as the sentence to the right of the arrow in (2a). As a result, the end product of the derivation may be a string with a mirror image pattern ($if_1 \ldots if_2 \ldots if_n \ldots then_n \ldots then_2 \ldots then_1$). This mirror image pattern cannot be generated by a finite-state grammar, since this device

[2] A point of clarification is in order at this point. Our discussion in this chapter focuses on narrow syntax, a core component of the model of grammar, but not the only component. In particular, the model includes interface components dealing with sound and meaning. For a fuller treatment, the question of the complexity of the grammar has to be answered separately for the grammar as a whole and for the individual components (including narrow syntax), with different answers forthcoming in each case. It is also an open question which phenomena are to be associated with which component of the grammar, with current proposals relocating seemingly narrow syntactic phenomena such as head movement and inflectional morphology to the interface component dealing with sound (e.g. Chomsky 2001). We abstract away from these questions and proceed on the understanding that narrow syntax involves nothing more than a sequence of operations Merge, joining elements from a predetermined set (the Numeration) into a hierarchical phrase structure.

computes a string strictly locally and thus cannot ensure an equal number of *if*s and *then*s (see Chomsky 1956 for a more rigorous presentation of the argument).

The relation between formal grammar and processing complexity had been addressed extensively by Chomsky and Miller's (1963) seminal work. By referring to the property of recursive self-embedding, they argued in favor of drawing a sharp distinction between processes at the level of performance and mechanisms at the level of formal grammar. As is well known, their observation was that multiple center-embedding leads to structures that cannot be produced or comprehended under normal on-line conditions, as illustrated by (3):

(3) The rat the cat the dog chased killed ate the malt.
(Chomsky and Miller 1963: 286)

Chomsky and Miller argued that the fact that such sentences are quite incomprehensible has no bearing on the desirability of generating them at the level of formal grammar, because, as Chomsky (1963: 327) pointed out by means of an analogy, "the inability of a person to multiply 18,674 times 26,521 in his head is no indication that he has failed to grasp the rules of multiplication." In other words, such structures are more complex than others due to performance factors that limit the realization of our grammatical competence. In response, however, Reich (1969) was the first to propose an FSG capable of generating sentences with degree-1 center-embedding but not center-embeddings of degree 2 or higher. Recently, approaches similar to those of Reich have been pursued in a connectionist setting by Christiansen and Chater (1999) and Christiansen and MacDonald (2009).[3] These accounts not only argue that natural languages are not of a PSG-type; they also claim that complexity measurement according to the Chomsky hierarchy in general is not motivated. They observe that self-embedding structures of a certain degree are not attested in linguistic performance and therefore argue that they should not be generable by the grammar.

The crucial distinction between these 'performance-oriented accounts' and the generative approach to complexity measurement is very clear. According to the generative perspective, the performance limitations on recursive self-embedding are captured by factors extrinsic to the competence grammar (such as memory overload induced by distance, cf. Gibson and Thomas 1999; Gibson 2000). In contrast, performance-oriented accounts such as usage-based approaches claim that "constraints on recursive regularities do not follow from extrinsic limitations on memory or processing; rather they arise from interactions between linguistic experience and architectural constraints on learning and processing [. . .] intrinsic to

[3] In section 7.4, we argue that center-embeddings can also be generated by a finite-state device. However, while connectionist approaches and their precursors do not assume a mentally represented grammar that allows unbounded recursion, we propose a competence grammar that allows generating infinite center-embeddings.

the system in which the knowledge of grammatical regularities is embedded" (Christiansen and MacDonald 2009: 127). In other words, while the generative approach postulates a competence grammar allowing unbounded recursion, the performance-oriented accounts deny the mental representation of infinite recursive structure and, thereby, try to nullify one of the axioms of modern linguistic theory: the grammar–performance distinction. A detailed discussion of these two positions would take us too far afield here (for a more general discussion, see Newmeyer 2003).

In this chapter, we follow recent work by Trotzke, Bader, and Frazier (2013) and Trotzke and Bader (2013), who present empirical evidence in favor of the grammar–performance distinction in the context of recursive self-embedding. Accordingly, in our view, the measurement of computational complexity, as represented by the Chomsky hierarchy, cannot be fruitfully connected to performance complexity, in keeping with the arguments of Chomsky and Miller (1963). Instead, the generative perspective on measuring linguistic complexity both abstracts away from linguistic variation and from processing complexity and focuses on basic formal notions of computational complexity. This is most clearly evidenced in the most recent version of generative grammar, namely the Minimalist Program (MP).

According to Chomsky (1995: 221), the MP is "a research program concerned with [. . .] determining the answers to [. . .] the question: 'How 'perfect' is language?'" In other words, the MP explores the hypothesis that language is a system that meets external constraints imposed by other cognitive components in the most 'elegant' (read: economical) way. Accordingly, as pointed out by Wilder and Gärtner (1997), within the MP, 'economy' is not only understood as a methodological postulate concerned with providing the 'simplest' description of a linguistic phenomenon. Rather, economy is also understood as referring to a property of language itself. Given this notion of the human language faculty, computational complexity arguments in terms of 'least effort' metrics play an important role in linguistic minimalism (e.g. Chomsky 1991; Collins 1997).

Let us briefly turn to basic computational aspects of minimalism in order to demonstrate in what sense they can be regarded as computationally 'optimal.' When we turn to computational science, two basic components of an algorithmic procedure must be distinguished: its time complexity and its space complexity (for a more extensive discussion of what follows, cf. Manber 1989; Mobbs 2008). While the number of operations required to perform a specific task constitutes the time complexity of an algorithm, the amount of working memory required for the performance of a task represents its space complexity. Accordingly, reducing both the time and the space complexity of running an algorithm results in more computational optimality. Now, how do minimalist conditions on derivations and representations correspond to these fundamental concepts of computational complexity theory?

To begin an answer to this question, let us look at the Extension Condition, proposed by Chomsky (1995). This condition on syntactic derivation holds that

"Merge always applies in the simplest possible form: at the root" (Chomsky 1995: 248), that is, there is only one possible site at which to extend a phrase marker. This condition thus minimizes complexity in accordance with fundamental complexity metrics of computational science. As Mobbs (2008: 29) points out, postulating "more than one possible site at which to Merge, it would be necessary to search for the appropriate site in each case, increasing the operational load [= the time complexity, AT/JWZ] on computation." The general constraint on syntactic derivations that ensures that the system meets such abstract complexity measures is the assumption that natural language syntax, as understood in the minimalist sense sketched above, should be operating in a maximally economical way.

Non-minimalist frameworks such as Simpler Syntax (Culicover and Jackendoff 2005, 2006) have also addressed the problem of the measurement of grammatical complexity. For example, this theory assumes that "the complexity of syntactic structure involves the extent to which constituents contain subconstituents, and the extent to which there is invisible structure" (Culicover and Jackendoff 2006: 414). In particular, this account attributes a higher syntactic complexity to 'mainstream generative grammar' (Culicover and Jackendoff's term), since mainstream approaches, as understood in the Simpler Syntax framework, operate with covert levels of representation like 'Deep Structure' and 'Logical Form' and with 'invisible' elements in the syntactic tree. These 'extra' levels and elements increase the representational complexity of syntactic structures, as Culicover and Jackendoff (2006: 413) briefly point out in the context of control constructions. According to their view, the situation that *drink* in *Ozzie tried not to drink* does have a subject does not necessitate an account that postulates 'hidden' syntactic representations like 'PRO.' In contrast, the interpretation of *Ozzie* as the 'drinker' can be formulated as a principle of semantic interpretation, external to the syntactic component.[4] So, according to Jackendoff (2008: 197), the main critique of minimalism refers to its complex representational format, since minimalism "requires null elements, a covert level of syntax, and particular hitches in the syntax that correlate in theoretically dubious fashion with the semantic peculiarities of the constructions in question."

In section 7.3, we show how the minimalist approach to these aspects of the representational format partly converges with the Simpler Syntax model. Since information structural properties like focus and topic are often accounted for in terms of an enriched syntactic representation and covert syntactic operations like LF movement (cf. Chomsky 1976 and later work), we will turn to this issue. In section 7.4, we discuss derivational complexity, i.e. the complexity of the structure building

[4] Interestingly, with regard to the very same phenomenon, recent minimalist literature argues in a similar vein. For example, Hornstein, Nunes, and Grohmann (2005: 54) replace the Theta-Criterion, which operates at Deep Structure, with a 'Theta-Role Assignment Principle' that applies at the semantic interface (LF).

process that generates syntactic representations, and argue for points of convergence with Construction Grammar approaches.

7.3 Representational complexity, cyclicity, and Simpler Syntax

In this section, we focus on representational complexity and its measurement in linguistic minimalism and in the Simpler Syntax framework. Advocates of the latter have claimed that "[t]he Minimalist Program [. . .] assumes that the structures and derivations of Principles and Parameters Theory are essentially correct" (Culicover and Jackendoff 2005: 88). They go on to argue that the elaborate and abstract structures of Principles and Parameters Theory are to be discarded in favor of considerable reduction of the representational format. In Simpler Syntax, this is achieved by relegating a substantial body of phenomena to 'interface rules.' In what follows, we demonstrate that proponents of a "constraint- and construction-based minimalism" (Jackendoff 2008: 222) do not fully acknowledge the recent shift from a representational to a strong derivational theory of linguistic structure within the MP. We argue that, given this shift, which involves minimizing the representational format of the computational system, the recent 'dynamic' approaches within the MP share basic assumptions with the perspective advocated under the Simpler Syntax hypothesis. To illustrate, we focus on an issue that has frequently been discussed with respect to representational complexity: the representation of information structural notions in syntactic structure. But before turning to this particular issue, we first point out the basic characteristics of the dynamic derivational model assumed in the MP, where cyclic (phase-based, layered) computations play a central role.

7.3.1 *Redefining derivations vs. representations*

The idea that the computation of a syntactic structure proceeds phase by phase (Chomsky 2000 and later work) has important consequences for the representational configuration, since the minimalist model of the grammar no longer defines a single point of 'Spell-Out' handing an entire syntactic structure to the interface components LF (the semantic component) and PF (the articulatory component). Instead, each derivation contains multiple points of interaction between the syntactic component (narrow syntax) and the interface components (LF and PF), the actual number depending on the number of phases. This dynamic interaction with the interfaces is also true of the model proposed in Zwart (2009a, section 7.4), where phases are redefined as derivation layers, that is, finite sequences of Merge yielding a substructure to be processed by the interface components and to be used in a further (sub) derivation. In a model with phases or derivation layers, the derivation of a sentence is a *system of derivations*, with multiple subderivations each feeding the interface components separately. So, while syntactic structures were hitherto considered to represent the whole sentence at some particular level of representation, it is now argued that derivations consist of stages in which only parts of these structures are represented at the interfaces. Put differently, "while there are still what might be called PF and LF components, there are no *levels* of PF and LF" (Lasnik 2005: 82, emphasis in the original). Consequently, the model of grammar resulting from these considerations can in effect be regarded as 'level-free,' as also pointed out by Boeckx (2006: 77).

The basic intuition behind models with multiple Spell-Out is that 'chunking' the derivation in subderivations leads to a reduction of computational complexity. As Miller (1956) had shown in the context of limitations on the amount of pieces of structure needed to be held in memory, "[b]y organizing the stimulus input simultaneously [. . .] into a sequence of chunks, we manage to break [. . .] this informational bottleneck" (Miller 1956: 95). Using phases or derivation layers in linguistic computation builds on this result. Otherwise put, and referring back to basic computational aspects sketched in section 7.2, chunking the derivation in phases or layers reduces the space complexity by reducing the amount of working memory required for running the algorithm (cf. Chesi 2007 and his specification of complexity reduction in phase theory).

Moreover, phase-based or layered derivations lead to a less redundant system by reducing the number of independent cycles in the computation. More concretely, in Principles and Parameters theory (Chomsky 1981), there were at least three relatively independent generative systems, all operating on the same domain, but separately. In Chomsky's (2004a: 151) words, "[t]here was one that formed d-structure by X-bar Theory, which is basically cyclic. There's the transformational cycle, which is mapping d-structure to s-structure. There's a covert transformational cycle, which is mapping s-structure to LF, with the same kinds of transformations and also cyclic." Chomsky claims that a phase-based derivation, with its cyclic transfer property, comes closer to the ideal of a single-cycle architecture.

One consequence of reducing the complexity of the model of the grammar in the context of levels of representations and covert syntactic operations remains to be explored. Given the dynamic interaction with the interfaces postulated in this 'strong' (i.e., in effect, 'level-free') derivational approach (cf. Brody 2002 for different incarnations and 'strengths' of derivational theories), it has been argued that this model allows for a direct interaction between PF and LF, that is, "PF has access to both, the syntactic derivation of the phase [. . .] and the semantic interpretation" (Winkler 2005: 24). As is well known, this direct interaction between the phonological and the semantic component is a crucial feature of Jackendoff's (1997 *et seq.*) Parallel Architecture, which is the model of grammar assumed in Simpler Syntax. This point of convergence between minimalism and Simpler Syntax has, to our knowledge, so far been mentioned only by Winkler (2005), who notes that the strong derivational model within minimalism "turns out to be conceptually closer to Jackendoff's [. . .] tripartite parallel model of grammar than might be recognized at first sight" (Winkler 2005: 231, n. 8). However, she does not elaborate on this point and leaves it to a short comment in a footnote.

7.3.2 *Reducing representational complexity: a minimalist perspective on the syntax-pragmatics interface*

Let us bring down the comparison of recent minimalism and the Parallel Architecture model to tractable size by focusing on the analysis of one specific phenomenon, namely the pragmatics of left-periphery-movement (LP-movement) in German, a topic that is described in information structural terms in the literature. Ray Jackendoff has repeatedly argued that, especially in the context of information structure, "there are aspects of semantics that have no impact on syntax but *do* have an effect on phonology" (Jackendoff 2003: 658, emphasis in the original). Accordingly, following Jackendoff's argument, cases like focus expressed through prosody seem to require a direct phonology–semantics interface, where "interface principles [. . .] map directly between a string of phonological words and a meaning" (Jackendoff and Wittenberg, this volume, chapter 4).

In an influential version of the standard generative framework, the 'cartographic program,' information structural properties of sentences are accounted for by encoding properties of information structure in the syntactic representation (e.g. Rizzi 1997). Thus, these accounts consider information structural properties to be represented in the syntax as formal categories actively determining the syntactic derivation. However, as cases like phonologically expressed focus demonstrate best, such a syntactic feature seems to provide no more than a device to pass information from semantics to phonology. Accordingly, Jackendoff (2002: 409) concludes that such a syntactic feature "is simply an artifact of syntactocentrism, the assumption that everything in meaning has to be derived from something generated in syntax." We now argue that the goal of reducing representational complexity in minimalism is consistent with Jackendoff's view. Before going into detail here, let us point out that we do not reject the *descriptive* advantages of the cartographic framework. Approaching syntactic structures (and especially the clausal left periphery) from a cartographic perspective has proven to be incredibly fruitful. Since proponents of this approach are committed, by and large, to a rigorous methodology of description, they can rely on a large amount of previous work and thereby also refine our picture of the overall syntactic structure of heretofore under-researched languages (for this point, cf. Trotzke 2012). However, in this chapter, we are concerned with the complexity of *narrow* syntax—a domain that refers to universal computational properties of natural languages and thereby contributes to the achievement of *explanatory* adequacy. These two goals, of course, complement each other; as Rizzi (2013b: 213) points out, "an accurate map of the sequence [of functional projections] is the essential point of departure for further study, including the search for further explanation" (cf. also Ramchand and Svenonius 2013 in this regard). However, in the domain of explanatory adequacy, it can be argued that even 'weak versions' of the cartographic account are incompatible with the dynamic approach to syntactic structure assumed in some implementations of the MP (cf. Zwart 2009b). In accordance with Newmeyer (2009: 131, emphasis in the original), we would therefore like to argue "that oceans of functional projections on the left periphery represent a singularly *unminimalist* framework for capturing [. . .] variation. Given that few languages manifest the proposed cartography in its fullness, there is no benefit to proposing that UG provides the set of projections and their ordering."

However, according to the proponents of the cartographic approach, and in contrast to Simpler Syntax and also to our view, there is no tension between enriched representations as proposed in cartography and minimalist ideas of reducing the syntactic computation to a minimum. In particular, as Rizzi (2004) argues, a crucial point of connection between the cartographic approach and the MP is the core idea of computational simplicity. The cartographic approach, according to Rizzi, contributes to this notion by decomposing functional projections into simple structural units. Thus, regarding complexity measurement, "[l]ocal simplicity is preserved by natural languages at the price of accepting a higher global complexity, through the proliferation of structural units" (Rizzi 2004: 8). However, cartographic approaches, with their enriched syntactic representations, are at root incompatible with the strongly derivational assumptions of standard minimalism. To be specific, proponents of the cartographic approach postulate functional heads in the left periphery possessing designated features for focus, topic, and other information structural constructs. According to Chomsky's (1995: 228) economy conditions, however, "any structure formed by the computation [. . .] is constituted of elements already present in the lexical items selected [. . .]; no new objects are added in the course of computation apart from rearrangements of lexical properties." In other words, this 'Inclusiveness Condition' implies that syntactic operations can refer only to lexical features. But of course, lexical items cannot be viewed as *inherently* possessing information structure properties. Consequently, such properties, as Neeleman and Szendrői (2004: 155) note, "must be inserted after an element has been taken from the lexicon," and thus the postulation of discourse-oriented features and the functional heads hosting them violates the Inclusiveness Condition.

Let us now turn to an empirical phenomenon in German that challenges cartographic accounts and their consequent higher global complexity. Recall that, according to Culicover and Jackendoff's (2006: 414) complexity measure, "the complexity of syntactic structure involves the extent to which constituents contain subconstituents, and the extent to which there is invisible structure." In what follows, we show how the German data can be accounted for from a minimalist perspective in a way that both reduces global complexity and is quite compatible with the Simpler Syntax approach.

In German, given certain pragmatic conditions, parts of idiomatic verb phrases can show up in the left periphery (for discussion of more phenomena that involve similar issues as discussed in this section, cf. Trotzke 2010). Consider the following example (for similar cases, cf. Fanselow 2004: 20):[5]

(4) den Löffel abgeben ('to die,' lit. 'the spoon pass')
[Den LÖffel]$_i$ hat er t$_i$ abgegeben.
the spoon has he passed
'He died.'

In (4), *den Löffel* receives pitch accent, indicating a contrastive interpretation. However, the preposed part *den Löffel* of the idiom *den Löffel abgeben* is meaningless in isolation (i.e. there is no set of alternatives to *den Löffel* in this particular context). Accordingly, fronting this element challenges the assumption of a dedicated syntactic position in the left periphery associated with focal or contrastive interpretation of the element that occupies this position, since, as Fanselow (2003: 33) puts it, "[i]t is difficult to accept the idea that a meaningless element can be interpreted as a focus or a topic [. . .] phrase." Thus, according to Fanselow's (2003) terminology, *den Löffel* is moved to the left periphery as a 'pars-pro-toto,' since the pragmatic interpretation involved is equivalent to preposing the whole idiomatic constituent, as shown in (5), where the whole predicate is fronted and interpreted contrastively.

(5) [Den LÖffel abgegeben]$_i$ hat er t$_i$.
the spoon passed has he
'He died (and did not survive).'

The fronting in (4), involving subextraction out of an idiomatic string, is problematic for any restrictive account of syntactic displacement. Even so, it appears to us that a cartographic approach, involving focus heads triggering LP-movement, does not provide a suitable model for the analysis of this phenomenon, since moving only a part of the focus hardly suffices to check a corresponding focus feature. In other words, we agree with Jackendoff's objection to enriching representational complexity by encoding information structure concepts in the narrow syntax.

In what follows, we propose a strongly-derivational minimalist approach that abstracts away from language-specific representations and also implements the idea that "a direct phonology-semantics interface [. . .] is attractive for the correlation between prosody and information structure" (Jackendoff 2003: 658). Thereby,

[5] In this and the following examples, capitals indicate relatively high pitch.

in accordance with Simpler Syntax, we aim at a considerable reduction of the syntactic representational format.

7.3.3 *A minimalist account of topic/focus interpretation*

We have seen that LP-movement to a topic or focus projection is not triggered by a topic or focus feature in a functional head in the left periphery, and that the topic or focus reading cannot be associated with elements at the point in the derivation where they are merged. Let us consider, then, an alternative approach in which topic or focus interpretation is an emerging feature, arising in the course of a derivation. In such a view, narrow syntax is oblivious to information structure; topic or focus readings are established only at the interfaces.

One way to approach the derivation of information structural interpretations is to say that Merge creates a structural dependency (or a set of such dependencies) that can be interpreted at the interface components in various ways (e.g. in terms of subject-predicate, topic-comment, or focus-ground oppositions, cf. Zwart 2009a: 170). From this perspective, the question of whether fronted elements appear in the left periphery via internal Merge (movement) or external Merge ('base-generation') is moot. What matters is that merging a left-peripheral element sets up a particular dependency allowing for a limited range of interpretations.[6] And just as Merge is not concerned with the notions of topic and focus, it is arguably not concerned with the corresponding prosodic effects. The fact that these effects nevertheless occur suggests that the interpretation established at LF (the semantic component) informs the relevant components of PF (the articulatory component) responsible for clausal prosody. This direct LF–PF interaction, however, does not in our opinion call for a radical revision of the model of grammar, as advocated by Jackendoff. Rather, the direct sound-meaning pairing results from a restrictive interpretation of the role of narrow syntax within the minimalist architecture.

To see that this interpretative approach to topic and focus interpretation is not unconstrained, consider once again the examples of LP-movement in German featuring the idiom *den Löffel abgeben* in (4) and (5). As we noted, LP-movement can front *den Löffel* alone, as in (4), or it can front the entire idiomatic expression *den Löffel abgegeben*, as in (5). The fact that *den Löffel abgeben* is a (noncompositional) idiom requires that the semantic component assembles the split parts of the idiom in (4), creating the idiomatic reading on the fly, as it were. We assume that this interpretive operation is a function of the composition of the two categories *den Löffel* and *hat er abgegeben* (cf. (4)), merged together at the final step of the narrow syntactic derivation. That is, for the correct interpretation of *abgegeben* in (4), the

[6] In what follows, we use the term 'LP-movement' to refer to a particular syntactic construction rather than to a syntactic operation.

interpretation needs to be supplied with *den Löffel*, which can be accomplished in the semantic component, once *den Löffel* has been merged to *hat er abgegeben* (again abstracting away from the question of whether *den Löffel* was moved or base-generated in left-peripheral position).

Let us make our analysis more concrete and state that the object in (6a), which is a constituent structure representation of (4) in the semantic component, has one realization in which the nonidiomatic material is ignored, as in (6b), so that just the idiomatic elements remain (cf. (6c)):

(6)	a.	(den Löffel)	(hat er abgegeben)
	b.	(den Löffel)	(~~hat er~~ abgegeben)
	c.	den Löffel	abgegeben

Informally, we can 'recognize' *den Löffel abgegeben* (6c) in *den Löffel hat er abgegeben* (6a), that is, we may take *den Löffel (. . .) abgegeben* (6a/c) as a token of the type *den Löffel abgeben* (cf. (4)). The operation in (6b) is allowed, we suggest, since the structural dependency between *den Löffel* and *abgegeben* is the same in the split and the unsplit idiom, with *den Löffel* merged with either *abgegeben* or a constituent containing *abgegeben*.

Observe now that splitting the idiom by fronting *abgegeben* is not possible:

(7)	# [Abgegeben]$_i$	hat	er	den	LÖffel	t_i
	passed	has	he	the	spoon	
	(intended: same as (4–5))					

From our perspective, this impossibility follows from the sensitivity of the semantic component to the reversed order of the idiom parts. To establish the idiom on the fly, *abgegeben* would need to be supplemented by *den Löffel* just as in (4), but the required structural dependency is not there, as *den Löffel* is not merged with a constituent containing *abgegeben*. In other words, the stripping procedure that worked well in (6), allowing the interpretation of (6b) as a token of the type (6c), now yields no result, as (8b) differs too much from (8c) for (8b) to be interpreted as a token of the type (8c):

(8)	a.	(abgegeben)	(hat er den Löffel)
	b.	(abgegeben)	(~~hat er~~ den Löffel)
	c.	den Löffel	abgegeben

On this analysis, the ungrammaticality of (7), where *abgegeben* is fronted alone, suggests a constraint not on LP-placement, but on idiom interpretation. Indeed, when idiomaticity is not a factor, constructions like (7) are possible, with the relevant context provided:

(9) [Abgewaschen]$_i$ hat er den LÖffel t_i (und abgetrocknet das Messer)
washed has he the spoon (and dried the knife)
'He washed the spoon.'

Returning to the main theme, we pointed out conceptual points of convergence between recent minimalism and the Simpler Syntax model in light of a derivational analysis of LP-movement in German. In particular, in contrast to approaches that are "forced to generate [. . .] sentences with a dummy syntactic element such as [+Focus], which serves only to correlate phonology and meaning" (Jackendoff 2003: 658), we demonstrated that information structural interpretation, from a minimalist point of view, emerges in the course of a derivation and thus is established only at the interfaces. In particular, the contrastive interpretation of the whole predicate *den Löffel abgeben* in both (4) and (5) is established at LF, which, in our model with direct LF–PF interaction, informs the relevant components of PF responsible for clausal prosody. This result is in line with the direct sound-meaning interaction postulated in Simpler Syntax, but it is also in accordance with other recent minimalist accounts that claim that "notions of information structure play no role in the functioning of syntax [. . .]. There is no reason left for coding information structure in the syntactic representation" (Fanselow and Lenertová 2011: 205). Our analysis, with its reduction of representational complexity, is thus preferable in the light of the representational complexity measure sketched by Culicover and Jackendoff (2005, 2006) that attributes a higher complexity to models operating with covert levels of representation and enriched syntactic trees. Note that excluding notions like focus or topic from the syntactic representation also avoids, to our mind, unwanted operations like moving constituents covertly to the relevant projection even if the focused or topical element stays in situ. Furthermore, it is a good exemplification of minimalism in general, which dictates "to examine every device [. . .] to determine to what extent it can be eliminated in favor of a principled account [. . .], going beyond explanatory adequacy" (Chomsky 2004b: 106). It is this 'going beyond explanatory adequacy' that can also be applied to derivational complexity, i.e. to the structure building process itself, as we show in the following section.

7.4 Derivational complexity, simplest Merge, and recursive layering

Recall from section 7.2 that early discussion of the complexity of a generative grammar (e.g. Chomsky 1956, 1957) established that the derivation of natural language recursive, self-embedding structures requires a formal system with the complexity of (at least) a context-free grammar. We see two developments in current linguistic minimalism that give rise to a new evaluation of the outcome of that discussion.

First, linguistic minimalism is characterized by a more articulated model of the grammar, where what takes place in the syntactic component ('narrow syntax') is reduced to the elementary process of combining elements ('Merge'). Other processes formerly considered to be part of syntax have been relegated to the interface components (including such hallmarks of transformational analysis as head movement and ellipsis, e.g. Chomsky 1995: 229). This new division of labor among the component parts of the grammar raises new issues as far as complexity measurement is concerned. In particular, the question of the complexity of grammar has to be answered separately for the grammar as a whole (narrow syntax, the interface components, and the interactions among them) and for the core component of narrow syntax, with different answers forthcoming in each case.

Second, linguistic minimalism is characterized by a more articulated concept of the derivation, i.e. the sequence of operations Merge turning a set of elements ('Numeration') into a hierarchical structure. In current minimalism, as mentioned above in section 7.3.1, such a derivation is taken to be punctuated, consisting of various phases (Chomsky 2001) or derivation layers (Zwart 2009a), each feeding the interface components independently of other phases/layers that are part of the same derivation. As argued in more detail below, this punctuated nature of the derivation calls for a reconsideration of what it is that makes a derivation recursive, thereby bearing on Chomsky's original argument for natural language grammars as being of the context-free type.

Let us begin the discussion of the complexity of grammar in linguistic minimalism with the hypothesis that narrow syntax is a maximally simple system containing only the structure building operation Merge. Merge selects two elements and combines them, yielding a set (Chomsky 2001: 3). Hierarchical structure is derived because the newly formed set is among the elements that Merge can select for its next operation. Since the output of one Merge operation can be the input for the next, Merge is taken to yield the property of recursion (we will modify this conception of recursion below). To facilitate a comparison between Merge and the rewrite rules of early generative grammar, it is necessary to consider the process from a top-down perspective, with the newly created set as the start symbol for that particular rewrite rule (to the left of the arrow), to be rewritten as the pair of elements merged. From the nature and number of those elements (terminal or nonterminal), it is possible to determine the complexity of the structure building process. Restricting ourselves to finite-state and context-free grammars, the type of rewrite rules to consider are:[7]

(10) a. finite-state: $A \rightarrow a \mid a\ B$
b. context-free: $A \rightarrow (a)^*\ (B)^*$

[7] In (10), ordinary characters are terminals and capitals nonterminals, | is the disjunction symbol, an* indicates an arbitrary number, and order is irrelevant.

The crucial difference between (10a) and (10b) appears to be the restrictions in terms of the number and nature of the elements yielded by the rewrite rules: in a finite-state grammar, the number of elements is at most two, of which at most one is a nonterminal, while in a context-free grammar, any number of terminals and nonterminals can appear. Note that we are not considering the *empirical* adequacy of each type of rewrite rule here, rather just the question of which type best captures the operation Merge, as currently understood. It would seem that the finite-state type (10a) comes closest, as the context-free type (10b) does not reflect a crucial property of Merge, namely that it combines two and only two elements. This suggests that the question of the complexity of the grammar, when applied to the component of narrow syntax, potentially yields a different answer ('finite-state') from the complexity question applied to the grammar as a whole, which includes narrow syntax, plus the interfaces, and the interaction among these components.[8]

There are, nonetheless, a couple of discrepancies between Merge and the finite-state rule type (10a). First, nothing in the rule seems to exclude the possibility that Merge can combine two nonterminals. Second, the finite-state rule may yield a single terminal, but Merge (by definition) cannot combine a single element (as Merge is defined as a process combining two elements). Perhaps both discrepancies disappear if a sister pair consisting of two phrases is disallowed because of labeling problems (Chomsky 2013) and if 'self-Merge' is what starts the derivation from a single terminal (Adger 2013: 19). However, it seems to us that there is a more natural way in which both discrepancies can be shown to be illusory.

First, we may simply assume that the first Merge operation in a derivation involves 'zero' as one of its terms (Zwart 2004; Fortuny 2008; De Belder and Van Craenenbroeck 2011), so that (10a) should be read as producing a terminal and either a non-terminal or nothing. Merge can produce a terminal and either a non-terminal or nothing, if we take Merge to be an operation that takes an element from some resource and adds it to the object under construction. The element taken from the resource is by definition a terminal, and the object under construction is typically more complex, being the result of earlier steps. This essentially iterative procedure yields binary branching, but since the object under construction is empty at the first step, the output of the first operation will be just a single terminal (Zwart 2004). This conception of Merge covers both parts of the disjunction in (10a), the rewrite rule yielding a terminal at the first step of the procedure, and a pair consisting of a terminal and a nonterminal after that. Similarly, in the top-down procedure proposed in Zwart (2009a), the final step in the procedure will involve just a single terminal, leaving only the empty set as what remains of the Numeration. Since the

[8] If so, sentences like (2a–b), which Chomsky argued show that natural language grammar is not context-free, should be reconsidered, for which see Zwart (2014) and below.

Numeration, in this system, is a nonterminal, each step yields a pair of a terminal and a nonterminal, except the last, which yields just a terminal (or a terminal and an empty set).

The discussion in the previous paragraph shows that slight modifications in the conception of Merge yield the result that Merge is essentially a finite-state type rewrite rule. Importantly, these modifications should not be seen as complications: whereas the standard conception of Merge stipulates that the number of elements manipulated should be two, these modified conceptions involve rules that manipulate just a single element at each step of the procedure (either adding it to the object under construction or splitting it off from the Numeration).

Second, the restriction that Merge, if it is to be of the type in (10a), never yields a pair of nonterminals, follows immediately once we understand that 'terminal' and 'nonterminal' should be defined relative to a derivation. This is where the punctuated nature of the derivation, mentioned above, becomes relevant. Assume as before that every derivation starts from a predetermined set of elements (the Numeration). Then a derivation maps the Numeration to a hierarchical structure. A terminal, then, is an element from the Numeration, whereas a nonterminal is essentially a stage in the derivation (e.g. the output of a step in the derivation). Since the Numeration may contain a phrase (construed in a previous derivation), a distinction needs to be made between the concepts of terminal/nonterminal on the one hand, and head/phrase on the other: A terminal can be either a head (X^0) or a phrase (XP). The finite-state rule (10a), then, while stating that a pair of nonterminals cannot be generated, does not state that a pair of phrases cannot be generated. What it does state is that one of the elements generated must be taken from the Numeration, its phrase structure status being irrelevant. Conversely, if an element of the Numeration is a phrase (the output of a separate derivation), merging it to the object under construction (yielding a pair of phrases) can be barred only by stipulation, and therefore should not be disallowed.

That derivations can be layered, with one subderivation feeding another, need not be stipulated. Indeed, it seems to be a feature of each grammatical component. For example, constructions manifesting derivational morphology, including compounds, are generated in separate, self-contained derivations, whose outputs may enter as atomic items in a syntactic derivation. Hence there is no reason to believe that this cyclic organization of the derivation should stop at the arbitrary boundary of 'words,' or that we could not have separate subderivations inside the syntactic component, one subderivation generating a phrase or a clause which shows up as a terminal in another. Essentially, this is the system of generalized transformations of early generative grammar (cf. Chomsky 1975 [1955]: 518). Arguably, the notion of *construction* as entertained in Construction Grammar (Goldberg 1995; Croft 2007) could be incorporated in the minimalist framework if we take constructions to invariably be the output of such a subderivation (Zwart 2009c).[9] Thus, the layered-derivation architecture contemplated here, to our mind, shares basic assumptions with the 'combinatorial interface rules' proposed by Jackendoff and Wittenberg (this volume, chapter 4). In particular, since Jackendoff and Wittenberg argue for uniform combinatorial operations regardless of "whether the constituent C is an utterance, a phrase, or a word, and whether its parts are phrases, words, or sub-word morphemes," their model shares our basic assumption that the opposition between words and phrases is artificial in the context of deriving syntactic objects.

The layered-derivation analysis might in fact be forced upon us by considerations of derivational simplicity. A grammar deriving a simple sentence like *The man left* from the Numeration in (11a) would not be able to proceed in a maximally simple fashion, merging a single element with each step of the derivation, as doing so would wrongly generate the constituent *man left* (illustrated for bottom-up unary Merge in (11b)). To get *the man* as a constituent in the output of the derivation, the Numeration would have to be as in (11c), where *the man* is the output of a separate derivation layer, and hence a phrasal terminal.

(11) a. { the, man, left }
b. step 1: merge *left* yielding **left**
step 2: merge *man* yielding **man left**
step 3: merge *the* yielding **the man left**
c. { [the man], left }

We refer to Zwart (2009a; 2011a) for a characterization of layered derivations and the idiosyncratic properties (including locality) that they yield. As stated above, accepting the possibility of layered derivations, terminals can be phrases, and the differences between Merge (in its simplest conception) and the finite-state rule in (10a) disappear.

We conclude, then, that Merge, as currently understood in linguistic minimalism, has the formal characteristics of a finite-state rewrite rule. It follows that if narrow syntax contains just Merge, a core component of the grammar may be characterized as having the minimal complexity of a finite-state grammar. Naturally, applying the question of complexity to the grammar as a whole yields a different answer. Nevertheless, we take the conclusion that narrow syntax is finite-state to be potentially important, especially if narrow syntax is to be equated with the faculty of language in the narrow sense, as argued by Hauser, Chomsky, and Fitch (2002).

[9] Not allowing elements in the Numeration to be phrases would increase the complexity of the grammar significantly, as the derivation would then have to involve subroutines, with the derivation feeding back into the Numeration (e.g. Bobaljik 1995) or involving parallel routes (e.g. Citko 2005). However, as pointed out in Zwart (2014), if the output of these subroutines or parallel routes showed idiosyncratic sound-meaning properties, the subroutines or parallel routes would have to be connected to the main derivation via the interfaces, and we would essentially be looking at layered derivations.

If there is any merit to our conclusion that narrow syntax is finite-state, we need to reconsider the argument from recursive self-embedding sentences of natural language, such as (2a–b), which it will be recalled was taken to prove that the complexity of grammar of natural languages has to be at least of the context-free type. We submit that the punctuated nature of the derivation, in particular its use of derivation layers, calls for a reconsideration of the question of how recursion is brought about.

In fact, a layered-derivation architecture is inherently recursive, as the output of one instance of a procedure *P* (a derivation layer) serves as input for another instance of *P* (Chomsky 1975 [1955]: 518; Zwart 2011b; Trotzke and Lahne 2011). A derivation as a whole, then, can be recursive even if its constituent subderivations are little finite-state grammars. The early discussions of the formal properties of natural language grammars were correct in stating that natural languages like English are not finite-state languages. However, it seems to us that this does not justify the conclusion that the rules of grammar (Merge in its simplest conception) are not finite-state. The complexity of the grammar as a whole is not a function of the nature of Merge, but of the recursive process of derivation layering. And the complexity of language is greatly increased by whatever takes place at the interfaces, which is irrelevant to the proper characterization of Merge. Accordingly, as for the complexity of the grammar as a whole, we acknowledge the reasoning behind the idea of "moving in the direction of construction-based grammar, which makes the 'interface' the heart of the entire grammar" (Goldberg 1996: 14).

Our discussion leads to the conclusion that the complexity of the grammar needed to derive recursive self-embedding sentences follows automatically from the punctuated nature of the derivation, i.e. from the assumption (inevitable, if we are correct) that derivations can be layered complexes of maximally simple subderivations. The arguments against grammar being finite-state, based on sentences like (2) and (3), then, are relevant to the complexity of the grammar as a whole, that is, narrow syntax in conjunction with the interface components and, crucially, the interaction among derivation layers, but not to the complexity of narrow syntax itself. It would take us too far afield to discuss the derivation of sentences like (2) here in any detail. However, a crucial property of these constructions may be pointed out, namely that they inevitably involve clausal embedding. If clausal embedding involves layered derivations (i.e. the interaction of subderivations, essentially in the form of generalized transformations), then constructions of the type in (2) are relevant to the question of the complexity of the grammar as a whole, not to the question of the complexity of each subderivation (i.e. each individual sequence of Merge operations). As before, the conclusion that English is not a finite-state language does not entail that narrow syntax is not finite-state.

7.5 Conclusion

In this chapter, we have discussed minimalist notions of formal complexity measurement and argued that both the derivational and the representational consequences of linguistic minimalism result in an approach that shares basic assumptions with alternative perspectives represented in this volume. After pointing out the distinction between complexity measurement at the level of performance and at the level of grammar in section 7.2, we sketched core minimalist ideas of complexity measurement. In section 7.3, given this general background, we showed that reduction in the domain of representational complexity results in points of convergence between minimalism and a "methodology of assuming as little syntactic structure as necessary" (Jackendoff and Wittenberg, this volume, chapter 4). However, while proponents of Simpler Syntax often call their approach a 'different sort of minimalism' by referring to the amount of representational levels and of covert operations in mainstream generative grammar, we showed, based on our analysis of one specific phenomenon from German, that complexity reduction in the sense of Simpler Syntax follows from the standard minimalist methodology. After discussing representational complexity in the context of left-periphery-movement in German, we turned to the issue of derivational complexity from the perspective of current linguistic minimalism. We argued that, as far as overall complexity is concerned, narrow syntax must be distinguished from the interface components, and one must keep in mind the punctuated nature of the derivation, which involves phases and/or derivation layering. In sum, narrow syntax can be captured by a finite-state device and, therefore, falls low on the Chomsky hierarchy. Furthermore, given that not only words, but also phrases and clauses can be atomic items in the course of a syntactic derivation, we concluded that the notion of 'construction,' as entertained in alternative frameworks, can be fruitfully implemented in a minimalist system.

8

Constructions, complexity, and word order variation

PETER W. CULICOVER

8.1 Introduction

This chapter is concerned with the possibility of accounting for word order and word order variation in terms of complexity.[1] I propose that it is useful to consider word order variation in terms of competing constructions, where other things being equal, the less complex construction is preferred by speakers. This view of variation presupposes that we have a way of measuring complexity. I suggest that both formal complexity and processing complexity play a role in driving change and variation. Formal complexity has to do with how much idiosyncrasy and irregularity has to be

[1] This chapter is an adaptation and extension of a number of topics discussed in Culicover 2013. It was made possible by the generous support of the Alexander von Humboldt Stiftung. I am grateful to the participants in the Conference on Formal Linguistics and the Measurement of Grammatical Complexity, University of Washington, March 23rd and 24th, 2012, and audiences at the Ohio State University, Humboldt University in Berlin, the University of Tübingen, the University of Wuppertal, and UC Davis for useful questions and comments. I thank also Erhard Hinrichs, Marga Reis, Andrzej Nowak, Susanne Winkler, Ray Jackendoff, Guido Seiler, Jack Hawkins and Martin Salzmann for their support and suggestions. I owe a particular debt to Laurel Preston and Fritz Newmeyer for their detailed and constructive editorial feedback. Any errors are my responsibility.

Since space is limited, I take the opportunity to cite here the work that has inspired much of my thinking about these issues.

- Markedness and optimality (Chomsky 1964, 1965)
- Processing complexity and dependency (Gibson 1998, 2000; Hawkins 1994, 2004; Haider 1995, 1997, 2000, 2003)
- Empirical work on verbal clusters and word order in Continental West Germanic [CWG] (Wurmbrand 2004, 2005; Haegeman and van Riemsdijk 1986; Zwart 1995, 1996; Bader and Schmid 2009a, 2009b, 2009c; Schmid 2005; Schmid and Vogel 2004; Vogel 2003, 2009; Bies 1996; Sapp 2005, 2006, 2007)
- Factors bearing on word order (Lötscher 1978; Wasow 1997, 2002, 2009; Hawkins 1982, 1994, 1999, 2001, 2004, 2006)
- Linearization and structure (Curry 1961; Dowty 1996; Kathol 2000)
- Social dynamics and epidemiology (Nowak and Latané 1994; Seiler 2008; Enfield 2008)
- Constructional approaches to grammar (Fillmore *et al.* 1988; Kay and Fillmore 1999)

expressed by a grammar, while processing complexity concerns the computational resources required by a language user to map between syntactic forms and conceptual structures. Focusing on English infinitival relative clauses and Continental West Germanic verb clusters, I suggest several ways in which complexity may be measured and how such complexity may contribute to language change and variation. I consider how complexity may actually arise in the course of change, and why it may persist even in the face of pressures to reduce it.

I originally became interested in the issue of constructional variation when I was working on the evolution of English *do*-support (Culicover 2008). One of the functions of *do*-support is to preserve the adjacency of the main verb and its complement, even in questions where there is inversion, as in (1).

(1) Did any of the students [pass the exam]?

In contrast, in a V2 language that lacks *do*-support, the verb often becomes separated from its complements, as in (2).

(2) *Old English*

ða	wendon	hi	me	heora	bec	to.
then	turned	they	me	their	backs	to

'Then they turned their backs to me.'
[Allen 1980: 287]

It is at least plausible that because *do*-support permits reduction of the dependency distance, it is less complex in some sense than the V2 alternative. In fact, Hawkins 2004 has proposed that the grammars of languages are organized in such a way as to minimize dependency distance.

However, assuming that in general grammars value simplicity and efficiency, explaining *do*-support in terms of the reduction of dependency distance raises an obvious puzzle: why don't all languages have *do*-support and eliminate the V2 alternative? That is, why are there languages like German and Dutch?

In the case of *do*-support it is plausible that there are in fact multiple dimensions of grammatical complexity that play a role in privileging each construction over the other. On the one hand, there is an advantage to introducing a construction such as *do*-support, which reduces dependency distance. On the other hand, there is a competing advantage to a V2 language, which preserves the form of the verb regardless of whether the verb is adjacent to its complements (as in (1)), or separated from its complements (as in (2)). The advantage afforded by the V2 language is arguably that the thematic structure governed by the verb is more readily identified when the main verb is inverted than when a dummy modal such as *do* is inverted. On this scenario, pressure to reduce complexity on one dimension may conflict directly with pressure to reduce complexity on another dimension. Languages or dialects may differ in that one pursues reduction of complexity on one dimension, while another pursues reduction of complexity on another dimension.

But then the question arises: why does one complexity factor win out in some languages, while another factor wins out in others? My suggestion is that the choice of which factor wins out when there are competing factors is not strictly a linguistic matter, but reflects social factors as well.

The structure of this chapter is as follows. §8.2 addresses the question: how do we move from intuitions about complexity on various dimensions to some concrete measurement of complexity? I distinguish two types of complexity, formal complexity and processing complexity, and discuss how they might be measured.

Section 8.3 focuses on a well-documented instance of word order variation, the Continental West Germanic (CWG) verb clusters. The simplest examples of these clusters consist of sequences of two verbs, in the order V_1-V_2 or V_2-V_1, where V_2 is the head of the complement of V_1 in a standard syntactic representation as in (3).

(3) $[_{VP}\ V_1\ [\ V_2 \ldots]]$

There are also three-verb clusters, which show all possible orderings of V_1, V_2, and V_3, V_3 being the most deeply embedded verb; some of these clusters are very well-attested, while others are very rare. The data shows that in many language varieties[2] it is not possible to say that all two- or three-verb clusters show a particular order. Which order is possible depends on the lexical subclasses of V_1, V_2, and V_3 that participate in each particular ordering. This sensitivity of order to lexical subclass suggests that word order is specified by CONSTRUCTIONS, that is, form–meaning correspondences. I discuss how to formulate such constructions in §8.4.

Proceeding from the assumption that alternative orders may emerge as a way of reducing complexity on different dimensions, in §8.5 I propose two dimensions of complexity that may account for the observed variation in CWG verb clusters.

The role of social factors in explaining the distribution of properties is taken up in §8.6. Given the properties and distributions of the CWG verb clusters, I offer the following scenario: social factors, such as network topology, geography, and different frequency distributions over speakers may favor one constructional alternative over another in different geographical areas, leading to variation. Contact leads to spread of properties, resulting in mixed varieties. I illustrate these points with a computational simulation. Section 8.7 concludes with a summary.

8.2 Measuring grammatical complexity

Two of the most prominent notions of grammatical complexity in the literature are what I refer to here as formal complexity and processing complexity.[3] Formal

[2] 'Language variety' is a cover term used in sociolinguistics for languages, dialects, registers, and so on.

[3] For other terminology and related discussion, see the papers in Miestamo (2008), as well as the chapters by Newmeyer and Preston and Moran and Blasi in this volume.

complexity has to do with the degree of generality of a grammatical description, in the spirit of markedness theory (Chomsky 1964, 1965; Battistella 1996). The intuition underlying this notion of complexity is that more general phenomena are simpler and more 'natural'. Irregularity, exceptionality and idiosyncrasy contribute to greater complexity, and, in principle at least, this complexity can be measured in terms of the number of terms, statements and length of statements in a description of the phenomena using some standard formal vocabulary.[4]

Processing complexity, on the other hand, has to do with the computational resources that are required by language users to map between a string of words and an interpretation. It is assumed in syntactic theory that processing complexity is not represented in the grammar per se (cf. e.g. Chomsky and Miller 1963). It has been proposed that differences in processing complexity may explain frequency differences in constructions with the same function, and, in the limit, the non-occurrence of a particular construction (Hawkins 1994, 2004).

For concreteness, I give one example of each type of complexity. I show in §8.2.1 how reduction in formal complexity may be appealed to in order to explain some instances of language change. In §8.2.2 I turn to processing complexity, and in particular, its role in determining the frequency of competing constructions in the linguistic experience of learners. In the discussion of CWG verb clusters to follow, I argue that since the formal complexity of the competing constructions is the same, the variation should be understood in terms of processing complexity.

8.2.1 *Formal complexity: an example*

Formal complexity can be illustrated by the English infinitival relative. Fillers in infinitival relatives must be a PP (4b), rather than an NP (4a).

(4) a. *the man who to talk to __
 b. the man to whom to talk __

In this way, they differ from infinitival questions, where the filler may be an NP.

(5) a. I wonder who to talk to __.
 b. I wonder to whom to talk __.

The grammaticality of (5a) suggests that the ungrammaticality of (4a) is an idiosyncrasy that must simply be stipulated in the grammar of English (Sag 1997; Culicover 2011).[5]

[4] See Chater and Vitány (2007) for discussion of the relationship between simplicity, description length, and language acquisition.

[5] For discussion of a number of other syntactic idiosyncrasies, see Culicover 1999, 2013. The impossibility of (4a) has been recognized in generative grammar for quite some time (cf. Emonds 1976: 192–195), but has not always been viewed as a gap in an otherwise regular pattern.

The presence of idiosyncrasy is somewhat puzzling on the view that languages seek to reduce formal complexity. One might think that all exceptions would be eroded away over time. English infinitival relatives offer an intriguing insight into how complexity as exemplified by (4a) may actually arise as a language changes in the direction of greater generality. In Old English (OE), infinitival relatives lacked relative pronouns entirely (Kemenade 1987: 151). They used either zero or the equivalent of *that*. When the relativized phrase was a constituent of PP, the clause was a zero-relative and the preposition was stranded, as illustrated in (6).

(6)	Drihten,	Þu	Þe	gecure	Þæt	fæt	[$_{PP}$ on *t*]	to eardienne
	Lord,	you	yourself	chose	that	vessel	in *t*	to live

[Blick[6] 157/ Kemenade 1987: 151]

But finite relatives had relative pronouns. The OE tensed relatives did not permit P-stranding with a clause-initial *wh*-; only clause-initial pied-piped relative PP was possible (Kemenade 1987: 152–153). It is therefore plausible to assume that infinitival *wh*-relatives are an innovation. The innovation was in the direction of increasing uniformity, by extending *wh*-relatives from the tensed to the infinitival cases. The scenario is illustrated in Figure 8.1. Stage 1 is the extension of *wh*-relatives to infinitivals, and Stage 2 is the extension of infinitival relatives to PPs.

The generalization began by extending relatives from zero-tensed to zero-infinitival. The next stage, still manifested by present-day English, extended infinitival relatives to clause-initial PP, leaving out clause-initial NP. The resulting form is an idiosyncrasy when interpreted with respect to tensed relatives, which allow all three types, as seen in Figure 8.1. In other words, the pressure to reduce formal complexity in the general domain of relative clauses resulted in an increase in formal complexity in the more restricted domain of infinitival relatives, marked by the stipulation in the grammar that the overt relative in an infinitival relative must be a PP.

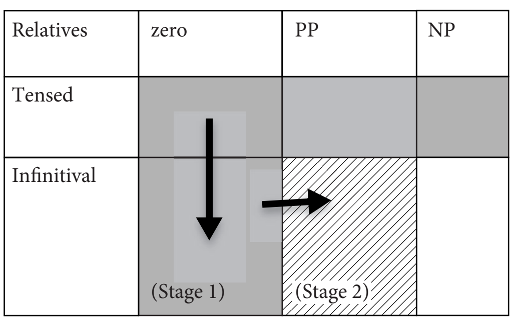


FIGURE 8.1 Evolution of relative clause types in English.

6 'Blick' refers to the Blickling Homilies, dated around the end of the tenth century (Westlake 1908).

A likely possibility, given *wh*-NP in initial position in infinitival questions (i.e. *wonder who to talk to*), is that the last cell will be filled in over time, and in fact isolated cases of this can be found, as shown in (7):

(7) a. Where do I find *the person who to* talk to about the quest?
http://answers.yahoo.com/question/index?qid=20100727153550AAPs5ed

b. In this case, I'll refer to the radio/club DJ as being *the person who to* target.
http://independentmusicstartup.com/440/how-to-get-a-tastemaker-to-take-your-music-to-the-next-level/

c. All requests for aid should include: . . . 3. The name and identity of the requesting person or the *person who to* contact upon arrival.
http://www.jdcap.org/SiteCollectionDocuments/EmergencyPlanExternal.pdf

d. That I needed the name of the Mortuary and number and *the person who to* speak to to verify if it was true . . .
http://www.prisontalk.com/forums/archive/index.php/t-488639.html

My native speaker judgment is that these examples are quite acceptable, in contrast with (4a), but only if the head noun is *person* (and not, for example, *man*). It is possible that they reflect the earliest stages in the elimination of the idiosyncrasy that constitutes formal complexity, namely, that the relative marker in an infinitival relative must be PP. Of course, it cannot be known at this point whether the change will generalize or simply die out.

8.2.2 *Processing complexity*

The second type of complexity is relevant to the processing of syntactic structures and assigning them interpretations. This type of complexity may be observed in terms of measures such as eye-tracking, self-paced reading, and reaction times (Gibson 1998, 2000; Levy 2005, 2008; Demberg and Keller 2008). For example, it is known that subject relatives (*the doctor that consulted the nurse*) require less processing (as measured by reading time) than object relatives (*the doctor that the nurse consulted*), although both types of relatives are fully grammatical. Self-embedding contributes further to processing complexity, so that sentences like (8c) are virtually impossible to comprehend.

(8) a. The doctor visited the patient.

b. The doctor that the nurse consulted visited the patient.

c. The doctor that the nurse that the hospital hired consulted visited the patient.

Hale (2001) proposes that the greater difficulty of processing configurations such as object relatives and self-embedding correlates with the frequency of occurrence of such configurations in corpora, and thus the probability of the configuration. The lower the probability of a particular configuration, the higher the 'surprisal' at the point at which the processor is forced to switch from the expected parse to an unexpected parse (Hale 2001, 2003; Levy 2005, 2008). In extreme cases, the sequence of words up to a particular point in the sentence points strongly to one continuation, but the actual continuation is different, resulting in a garden-path effect (Frazier 1987).

The question then arises, what makes one construction less frequent than another? In some cases, at least, the answer may be that the less frequent construction is more complex in terms of production. Complexity has been attributed to the maintenance of representations in memory (Kluender 1998), the maintenance of reference and the cost of the operations that build structure (Gibson 2000), and the cost of backtracking and repair (Frazier 1987). On the assumption that speakers avoid the more complex in favor of the less complex (Hawkins 1994, 2004), other things being equal, the frequency of more complex constructions in the experience of learners will be lower. Hence there is a loop that links processing complexity for the speaker to frequency to processing complexity for the hearer.

In sum, there is a plausible case to be made for relating processing complexity, corpus frequency and acceptability, although the precise details of the relationship remain to be worked out. In the next section I summarize the data from CWG verb clusters, as a step towards outlining a scenario for understanding variation in terms of processing complexity.

8.3 CWG verb clusters

Verb clusters (exemplified in English by *will sing, has sung, try to find,* etc.) are useful for exploring complexity and variation for two reasons. First, they show a substantial amount of variation among forms that have the same interpretation, and can therefore be compared directly in terms of formal and processing complexity. Second, there is a wealth of empirical data about the distribution of various types of verb clusters over the Continental West Germanic area, so that we can seek correlations among various types of clusters. This section summarizes the basic types of clusters and suggests that they should be analyzed as individuated constructions. The next section addresses how a constructional analysis of verb clusters might be formulated.

Verb clusters are sequences of modals (MOD), *have-* and *be*-type auxiliaries (AUX), and various control verbs (V). Typical two-verb clusters are given in (9), using Standard German forms for the sake of illustration. The verb order is shown in parentheses, where 1 is the highest verb (i.e. the one closest to the root) in a standard syntactic analysis, 2 the next highest, and so on (cf. (3)). So in (9a), for example, 1 is *kann* and 2 is *singen.*

(9) Maria glaubt, daß
Maria believes that

a. sie die Arie singen kann. (2–1)
she the aria sing can
'... she can sing the aria.'

b. sie die Arie kann singen. (1–2)

The order 2–1 of (9a) is found in Standard German; the 1–2 order of (9b) is found in a number of Dutch dialects, Swiss German dialects and other CWG varieties. I suggest in §8.5 that 1–2 is a response to pressure to order scope taking elements, such as AUX and MOD, before their complements, while 2–1 is a response to pressure to position heads as close as possible to their dependents.

The 3-verb clusters found in CWG are shown in (10), again using standard forms. 'Rare' means that the particular ordering is found only in a few varieties, and is highly lexically restricted.

(10) a. sie Peter die Arie singen hören wird. (3–2–1)
she Peter the aria sing hear will
'... she will hear Peter sing the aria.'

b. sie Peter die Arie hören singen wird (2–3–1) [rare]

c. sie Peter die Arie wird hören singen (1–2–3)

d. sie Peter die Arie wird singen hören (1–3–2)

e. sie Peter die Arie singen wird hören (3–1–2)

f. sie Peter die Arie hören wird singen (2–1–3) [rare]

An important property of CWG verb clusters is that some of them are very frequent, others are very rare, and there are apparent correlations between two-verb and three-verb clusters. That is, certain three-verb clusters occur only in varieties with 1–2 in two-verb clusters, while others occur only in varieties with 2–1. Many varieties appear to permit more than one order, both in two- and three-verb clusters.

Figure 8.2 shows the distribution of two-verb clusters in nine different varieties (data are from Sapp 2011: 108; the headings are renamed for consistency).

Sapp highlights the fact that for each variety and for each subtype of cluster, there are different preferences for word orders (not shown in Figure 8.2). For example, he cites Patocka (1997) as observing that in the Austrian dialects, the passive with *werden*-V.PAST.PRT is uniformly 2–1, while *haben*-V.PAST.PRT permits both orders in some cases.

Along similar lines, Dubenion-Smith (2010:112) found the judgments in Figure 8.3 in an analysis of spoken West Central German[7] from the *Datenbank Gesprochenes*

[7] The area bounded roughly by Karlsruhe to the South, Darmstadt to the East, Kassel to the North, and Aachen to the West.

Dialect	MOD V	AUX V
Standard German	2–1	2–1
German & Austrian dialects (Wurmbrand)	2–1	2–1
S and W Austria	1–2 (2–1)	1–2 2–1
N Austria	2–1	2–1
E Austria	2–1	1–2 2–1
Bavarian	2–1	2–1 (1–2)
Swabian	2–1	2–1 (1–2)
Alsatian	2–1 (1–2)	2–1
Swiss	1–2 (2–1)	2–1 (1–2)

FIGURE 8.2 Word order in two-verb clusters (Sapp 2011: 108). Reprinted by permission of John Benjamins.

Deutsch compiled between 1955 and 1970. Each cell shows the number of tokens and percentage of the total number of tokens of each ordering for verb clusters containing verbs of particular classes, e.g. past participle V and AUX=*haben* 'have', *sein* 'be', *hätten* 'have.SUBJUNCTIVE', *wären* 'be.SUBJUNCTIVE', or *werden* 'Passive be', in the first row, V MOD in the second row, and so on. The data show that both 1–2 and 2–1 are possible for many two-verb clusters, and that different combinations of verb classes occur with different frequencies. (I have changed some of Dubenion-Smith's notations for the sake of uniformity.)

Dubenion-Smith's more detailed figures, not reproduced here, show that for V AUX, V MOD, V *tun* and V_2 V_1, both of the orders 1–2 and 2–1 are attested, with a strong preference for 2–1. Dubenion-Smith takes the useful step of breaking down the data into specific subtypes of two-verb clusters, which allows the variability

Syntagm (group)	2–1	1–2	Total tokens
V AUX (perfect, subjunctive, passive)	887 (92.4%)	73 (7.6%)	960 (72.4%)
V MOD	227 (74.9%)	76 (25.1%)	303 (22.9%)
V *tun* 'do'	27 (96.4%)	1 (3.6%)	28 (2.1%)
V_2V_1	13 (59.1%)	9 (40.9%)	22 (1.7%)
V *kriegen* 'get'	8 (100%)	0 (0%)	8 (0.6%)
V *lassen* 'let/make'	4 (100%)	0 (0%)	4 (0.3%)
MOD AUX	1 (100%)	0 (0%)	1 (0.1%)
Totals	1,167 (88%)	159 (12%)	1,326 (100%)

FIGURE 8.3 Two-verb clusters in spoken West Central German subordinate clauses (Dubenion-Smith 2010: 112). Reprinted by permission of Cambridge University Press.

to show through clearly. While the totals show that both orders are possible, the breakdown shows that the 1–2 order is more common for MOD V (as a percentage of the total cases) than it is for the other subtypes.

Similar variability is found in three-verb clusters. Consider the data in Figure 8.4 (from Schmid 2005) showing possible three-verb clusters in Zürich German, broken down by the type of verb that occupies the V2 slot. The data show that which three-verb clusters are possible in this variety depends on the verb types of V_1 and V_2. (The verb types are exemplified by *lassen* 'let/make' (Causative), *müssen* 'must' (Modal), *sehen* 'see' (Perception Verb), *helfen* 'help' (Benefactive), *bleiben* 'stay' (Durative), *beginnen* 'begin' (Inchoative), and *versuchen* 'try' (Control Verb).)

V_2	AUX_1 V_2[PAST.PRT] V_3	AUX_1 V_2[INF] V_3 (IPP)	FUT_1 V_2 V_3
Causative	*	3–2–1, 1–2–3, 1–3–2	3–2–1, 1–2–3, 1–3–2
Modal	*	?3–2–1, 1–2–3, 1–3–2	?3–2–1, 1–2–3, 1–3–2
Perception verb	3–2–1, ?1–2–3, 2–1–3	?2–3–1, 1–2–3	3–2–1, 1–2–3, 1–3–2
Benefactive	3–2–1, 2–3–1, 1–2–3, 1–3–2, 2–1–3	2–3–1, 1–2–3, 2–1–3	3–2–1, 2–3–1, 1–2–3, 1–3–2
Durative	3–2–1	*	3–2–1, 1–3–2
Inchoative	2–3–1, 2–1–3	*	2–3–1, 1–2–3, 2–1–3
Control verb	3–2–1, 1–2–3, 2–1–3	*	3–2–1, 1–2–3, 1–3–2, 2–1–3

FIGURE 8.4 Overview of verb order patterns in Zürich German (adapted from Schmid 2005: 76). Reprinted by permission of John Benjamins.

According to Schmid, "Zürich German shows the largest variation of verbal order patterns of all languages [. . .] With the exception of order 3–1–2 (only possible with a special stress pattern), all logically possible patterns are confirmed by my informant."[8] But what her summary also shows is that different orders are attested with different verb classes.

Data such as these show that it is not possible to say that two- and three-verb clusters in a particular language variety have a particular ordering. In some language varieties, particular orderings occur only with particular classes of lexical items. In some language varieties, more than one ordering occurs, but one order occurs significantly more frequently than the other.

The variety of attested orderings raises a number of questions: how do we account for the possible orderings in each variety? Specifically, is there a derivational account that explains the observed orderings in terms of an underlying 'canonical' order? Or

[8] However, according to another native speaker informant (Martin Salzmann, p.c.), not all of these orders are actually active in this dialect. It is entirely plausible, although of course difficult to confirm, that the judgments of Schmid's consultant were influenced by exposure to standard German and to other dialects of Swiss German.

must the various orderings be treated as distinct, but related, constructions? Why are some orderings more frequent than others? Does either formal or processing complexity have anything to do with these phenomena? If the rare orderings are more complex in some sense than the more common ones, then why have they not been completely supplanted by the less complex orderings? How are clusters properly integrated into grammatical descriptions in terms of syntax and semantics? That is, what is the relationship between the structure of a verb cluster and its interpretation? I propose some answers to these questions in the following sections.

8.4 Constructions

There are three properties of verb clusters that have to be accounted for. First, clusters with the same elements but different orders have the same interpretation. Second, all of the possible orders of clusters composed of the same elements occur in at least one language variety. And third, the clusters 1–2–3 and 3–2–1 are significantly more frequent than the others.

In this section I propose a constructional analysis, in which the verb clusters are individuated unheaded phrases. That is, they are constituents, members of a sui generis category, call it VerbCluster (VC).[9] Individual variants of a word order (e.g. AUX_1-V_2 vs. V_2-AUX_1) are distinct but related constructions, that is, they are form–meaning correspondences (in the sense of Culicover and Jackendoff 2005). On this view, a construction consists of a pairing of a syntactic configuration, and a conceptual structure. Crucially, the related constructions express the same correspondence between form (syntactic structure) and meaning (conceptual structure).

In order to show the explanatory power of the constructional account, I first outline the more traditional derivational approach, in which a verb cluster with a particular word order is derived from some canonical underlying structure by movement and adjunction (Wurmbrand 2004, 2005).[10] The synonymy of clusters with different orders is accounted for by deriving them from the same underlying structure, which corresponds to a particular interpretation.

Consider, for example, *das Buch lesen$_2$ kann$_1$* 'can read the book', with the structure in (11a), and the alternative *das Buch kann$_1$ lesen$_2$*, with the structure in

[9] For extensive arguments that syntactic theory should allow for sui generis constructions that exist in a language alongside more general constructions, see my book Culicover (1999). In Culicover (2013), I discuss the nature of the formal relationship between the more specific and the more general constructions in a language.

[10] I do not pursue such an approach in detail here, in part because of Wurmbrand's demonstration that there is no evidence that favors one candidate underlying order from another, in part because it is not clear how to account for the variability of the sort noted in the preceding section in terms of movement and adjunction, and in part because it does not offer a natural way to account for the relative rarity of some orderings versus others.

(11b), derived by movement from (11a). The corresponding conceptual structure for both expressions is (11c), where **READ**(AGENT:X, THEME:BOOK) corresponds to *das Buch lesen* and **ABLE** corresponds to *kann*.

(11) a. $[_{VP}\ [_{VP}\ [_{NP}$ das Buch] lesen_2] kann_1]
b. $[_{VP}\ [_{VP}\ [_{NP}$ das Buch] t_2] kann_1+lesen_2]
c. **ABLE**(**READ**(AGENT:X,THEME:BOOK))

Since the function **ABLE** takes as its argument the relation **READ**(AGENT:X, THEME: BOOK), it is customary to say that the modal scopes over its complement in both (11a) and (11b).

Different orders are accounted for on such a derivational account by stipulating that certain heads adjoin to the left or to the right of higher heads, along the lines of (11b). More complex derivations are required for three-verb clusters; a few are illustrated in (12), assuming underlying 3–2–1.

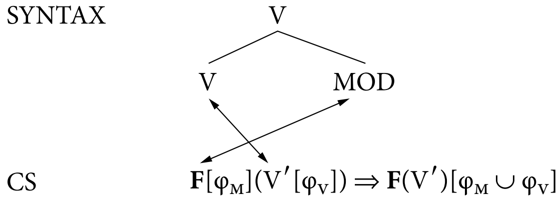
(12) a. $[_{VP}\ [_{VP}\ [_{VP} \ldots V_3\ ]\ V_2\ ]\ V_1\ ] \Rightarrow [_{VP}\ [_{VP}\ [_{VP} \ldots t_3\ ]\ V_2{+}V_3\ ]\ V_1\ ]$ (2–3–1)
b. $[_{VP}\ [_{VP}\ [_{VP} \ldots V_3\ ]\ V_2\ ]\ V_1\ ] \Rightarrow [_{VP}\ [_{VP}\ [_{VP} \ldots t_3\ ]\ V_2{+}V_3\ ]\ V_1\ ]$
$\Rightarrow [_{VP}\ [_{VP}\ [_{VP} \ldots t_3\ ]\ t_{2+3}\ ]\ V_1{+}V_2{+}V_3\ ]$ (1–2–3)
c. $[_{VP}\ [_{VP}\ [_{VP} \ldots V_3\ ]\ V_2\ ]\ V_1\ ] \Rightarrow [_{VP}\ [_{VP}\ [_{VP} \ldots V_3\ ]\ t_2\ ]\ V_1{+}V_2\ ]$ (3–1–2)

As far as I know there is no natural mechanism intrinsic to the grammar in such an account to explain why some orders are very frequent and others are not. One might surmise that the number of adjunctions would correspond in some sense to the complexity or markedness of a particular cluster. The application of two adjunctions shown in (12b) would then predict that 1–2–3 is very rare or possibly non-existent. However, the 1–2–3 and 3–2–1 clusters are the most common.[11]

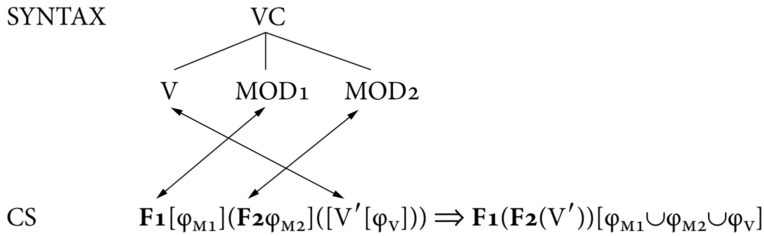
Let us consider now the constructional account. The problem of how to interpret the clusters is dealt with by unification, along the lines proposed by Bouma and van Noord (1998). In unification, the features of the constituents that form a phrase are assigned to the phrase itself. The semantics do not require hierarchical branching structure in the syntax, because the scope relation is established in the correspondence with conceptual structure: the representation corresponding to one verb takes the representation headed by the next as its argument. The selectional properties of the verbs are unified, as shown in (13) for a two- and in (14) for a three-verb cluster with two modals (e.g. $\textit{lesen}_3$ $\textit{können}_2$ $\textit{wird}_1$ '(s/he) will be able to read').

[11] If the underlying order were 1–2–3, then the incorrect prediction would be that 3–2–1 is rare.

(13)



(14)



So, following (13), the interpretation of [*lesen₂ kann₁*] will be (15), where **ABLE** corresponds to **F**, φ_M represents the selectional properties of *kann*, **READ** corresponds to **V′**, and φ_V represents the selectional properties of **READ**, here, the argument structure (AGENT:X,THEME:Y).

(15) **ABLE**(**READ**)[AGENT:X,THEME:Y]

Similarly, the interpretation of (16a) will be (16b).

(16) a. lesen$_3$ können$_2$ wird$_1$
read.INF can.INF will
'will be able to read'
b. **FUT**(**ABLE**(**READ**))[AGENT:X,THEME:Y]

In addition to (13) and (14) there exist constructions for each of the other possible orders, each of which has the same interpretation. For example, the construction for MOD$_1$-V$_2$, e.g. *kann$_1$ lesen$_2$* 'can read', will show the same correspondence with **ABLE** (**READ**) as *lesen$_2$ kann$_1$*, and the construction for MOD$_1$-MOD$_2$-V$_3$, e.g. *wird$_1$ können$_2$ lesen$_3$* will show the same correspondence with **FUT**(**ABLE**(**READ**)) as *lesen$_3$ können$_2$ wird$_1$*.

Consider now the fact that that the constructional approach presupposes that there are as many individuated constructions as there are distinct word orders with the same interpretation. From the perspective of description this is a virtue, since in fact all orders are attested. Moreover, the existence of all of these constructions does not render the constructional account overly complex, if we incorporate into our syntactic theory the further assumption that all orderings of the daughters of a phrase in a construction are available in a grammar, if at least one ordering of the daughters of that phrase is. In this respect I am drawing on an intuition that has been implemented in a variety of ways in the theoretical literature, most importantly in optimality theory (see for example Schmid and Vogel 2004), as well as in varieties of categorial grammar (Muskens 2007).

Given this assumption, the existence of a construction where V_2-V_1 corresponds to a specific interpretation immediately entails that V_1-V_2 is possible with the same interpretation. I assume, however, that whether a given construction is actually used or accepted by native speakers depends in part on the frequency with which they encounter instances of that construction, in part on the formal complexity of the construction with respect to the grammar as a whole (as discussed in §8.2.1), and in part on the processing complexity of the construction with respect to the available alternatives. The assumption that alternative orders of phrasal constituents are in principle generally available is critical, because it allows for an explanation of the spontaneous emergence in a variety of a construction as well as the growth of a competing construction under the influence of contact. Of course, not all orders are equally likely to occur, in part for reasons of complexity, as I discuss below.

The principled availability of all orderings of a construction also permits an explanation of why the 1–2–3 and 3–2–1 three-verb clusters are more frequent than the others, and why 2–3–1 and 2–1–3 are quite rare. 1–2–3 is optimal with respect to one of the complexity biases discussed in the next section, and 3–2–1 is optimal with respect to the other. As noted above, this asymmetry is not accounted for in a derivational account. It is in fact stipulated in OT accounts of CWG verb clusters (Bader and Schmid 2009a,b,c; Schmid and Vogel 2004), but not explained in terms of complexity. The constructional account I propose accounts for the relative frequency of the various orders in terms of complexity and the dynamics of the social network, as discussed in the next two sections.

8.5 The role of complexity biases in accounting for change and variation

Section 8.3 summarized data showing that the ordering in verb clusters is variable, across and within varieties, and even within speakers. I suggest in what follows that 1–2 is a response to pressure to order scope taking elements, such as AUX and MOD, before their complements, while 2–1 is a response to pressure to position heads as close as possible to their dependents. Why multiple orders can be simultaneously maintained is a matter that I take up in §8.6.

To get an insight into the biases that affect word order in verb clusters, consider the following quotation from Haider. Haider (2003: 28) argues that because of 'parser unfriendliness' there must be a constraint that requires branching nodes to follow the heads, thereby ruling out the left branching structure. Instead, there must be a cluster in V-final structures. (See also Dubenion-Smith 2010 for a similar idea.)

In a right branching structure, the parser can unambiguously identify the top-most node of the projection after encountering the first element of the projection, V1. [The left branching structure], however, is not parser-friendly. The parser would have to guess how many brackets there might be, because their number—or in other words, the depth of embedding of the left-most element—depends on the number of verbs to come. General top-down information on the possible structure of a VP will not help guessing, because the number of auxiliaries is not context dependent. (Haider 2003: 8)

A slightly different way to put Haider's observation is that there is nothing formally wrong with 3–2–1 order, per se, but 1–2–3 order is computationally less complex with respect to a particular aspect of processing. The tense or modal operator expressed by the first verb must take scope over the VP that is its argument, as discussed in §8.4. Hence, it appears that something along the lines of the following is correct (see Kroch 1974, for example).

(17) *Scope-Order principle*
The preferred scope ordering of operators corresponds to the left-to-right ordering of the phrases in the surface structure of the sentence.

This principle suggests the following scope bias.

(18) *Scope bias*
Alignment of scope with linear order facilitates one aspect of the computation of scope in the CS representation.

With this in mind, consider what steps are required in processing a 2–1 order, exemplified in (19).

(19) ...daß sie das Buch lesen$_2$ will$_1$.
that she the book to-read wants
'...that she wants to read the book.'

For simplicity, let us suppose that there is a certain cost to processing a verb that selects a VP complement that precedes it, and that the costs associated with each such verb are equal; let this measure be **c**, which I will call the 'scope bias'. This cost occurs because when the preceding verb is encountered, it is not known precisely what the following verb will be, or even whether there will be such a verb in some cases, as Haider suggests. Since we don't know exactly how the parsing proceeds, I assume that the cost of each basic operation is also **c**, to simplify the exposition and to make it simpler to compare alternative ways of resolving parsing conflicts and uncertainties.

At the point at which the word *lesen* 'read' is processed, the partial representation is (20).

(20) **READ**(PRO, BOOK)

A potential complication is that *sie* is singular and the singular form of *lesen* is *liest,* so *sie* is not the subject of *lesen* but of some other verb that follows *lesen*. However, *sie* has the same form as *Sie* 'they, you (polite)', and the infinitival form and the 3.PL inflected form are the same. So at the point of *lesen*, there is no way to predict that there will be another verb, as Haider points out. Either we hold *lesen* in memory as an ambiguous form, or we decide that it agrees with *Sie.*[12]

In either case there is additional processing at *will* 'want', which is ambiguously 1|3.SG. If the processor refuses to make a commitment regarding *lesen*, then it has to hold the form in memory, a cost of **c**, by assumption. If the processor decides that *lesen* was 3.PL, then it has to change the interpretation from *Sie* 'they, you (polite)' to *sie* 'she/her', which also costs **c**. In either case, the result is (21), with a marginal processing cost of **c**, also by assumption.

(21) **WANT**(PRO[3.FEMALE]$^{\alpha}$, **READ**(α,BOOK))

Suppose now that the order is 1–2. In this case there is no marginal cost to processing the first verb in the cluster, because of its position in the sequence. Processing ... *daß sie das Buch* is the same in both cases. Then the processor gets to *will*, which shows that the subject is the 3.SG *sie* and not 3.PL *Sie,* which takes a different form of the verb (*wöllen*).

(22) **WANT**(PRO[3.FEMALE]$^{\alpha}$, **F**(α,BOOK))

And when the processor encounters *lesen*, all that is required is to fill in the value of **F** as **READ** and link the object thematic role to BOOK.

Processing the order 2–1 is thus apparently more complex than processing the order 1–2 by **c**. So, one might ask, why aren't all of the dialects 1–2(–3) (not to mention all languages)? The answer, as suggested earlier, is that there is also a dependency bias that favors 2–1 and 3–2–1. I state it here as follows, paraphrasing Hawkins' (1994, 2004) formulation in terms of syntactic domains—the syntactic domain of a head H and its dependent D is the smallest part of the tree that dominates both H and D.

(23) *Dependency bias*
The preferred ordering of a head and its dependents is the order that permits the minimal syntactic domain that contains them.

As discussed at length by Hawkins (1994, 2004), languages tend to favor linearizations in which the subcategorized arguments of a head and closely related adjuncts

[12] If the first verb is a past participle, such as *gelesen* 'read.PAST.PRT', it is very likely that the following verb is a form of the auxiliary verb *haben* 'to have', but the exact form of this verb cannot be predicted. It could be infinitival *haben* or finite (present or past tense), and in the latter case, it is inflected for person and number.

are as close to the head as possible, thereby reducing the domain in which the dependency can be computed by the processor. Hawkins' reasoning is that the computational cost of constructing the CS representation, which is correlated with memory load, is greater when the heads and dependents are more distant from one another. A similar measure of complexity is argued for by Gibson (1998, 2000) and Grodner and Gibson (2005).

Consider again the processing cost of 1–2 versus 2–1. Let us assume that holding an argument or an adjunct in memory until its head is encountered incurs a cost of **c**′; this is the dependency bias. In the 2–1 variant of our example (19), there is no cost, since the direct object *das Buch* is adjacent to the verb *lesen*. But in the 1–2 variant, the verb *will* intervenes:

(24) ... daß sie das Buch | will lesen.

The processor may hypothesize at the point that *das Buch* is encountered that it is an argument of some verb, probably a direct object, and so corresponds to **F**(α,BOOK). But just as holding **READ**(α,BOOK) until *will* is processed incurs a cost when the order is 2–1, so does holding **F**(α,BOOK) until *lesen* is processed in the 1–2 cluster. Hence with respect to processing the dependency, processing 1–2 is more costly than processing 2–1 by **c**′. And when we deal with more complex VPs, the processing costs are correspondingly greater for both orders.

Order 1–2 is thus favored by the scope bias, and 2–1 by the dependency bias. Competing biases do not necessarily cancel one another. The constructions that they apply to are both active in the population of speakers, and in principle one or the other can acquire an advantage through an accidental property of the topology of the network. That is, a critical mass of speakers can shift to one order for non-linguistic reasons, making it the only order or the preferred order.[13]

Crucially, these two biases are in competition only in languages with VP-final clusters. In a language like English, where V is initial in VP, the 1–2(–3) cluster is optimal with respect to both the scope bias and the dependency bias. In languages with VP-final clusters, however, alternative orders may be seen as responding to one or the other of these biases. (25) illustrates, where X represents the complements of the main verb.

[13] What I am suggesting here is similar to, but different from, the notion of 'competing motivations', which as I understand it typically attributes different outcomes of language change to competition between formal and functional pressures, and between various functional pressures. In the case under discussion, I am proposing that different contributors to processing complexity are in competition and produce different outcomes in different social contexts. For discussion see Newmeyer (2005) and references cited there.

(25) X 3–2–1 ⇒ X **1–3–2** [scope bias]
X 1–2–3 ⇒ X **2–1–3** [dependency bias]
X 1–2–3 ⇒ X **2–3**–1 [weak dependency bias, assumes unitary 2–3; cf. Haider 2003]
X 1–2–3 ⇒ X **3**–1–2 [dependency bias]

All of these orders are attested, as shown below. 1–3–2 is the order found in the Infinitivo Pro Participio construction AUX-V-MOD, where MOD has the infinitival rather than the participial inflection, as in (26).

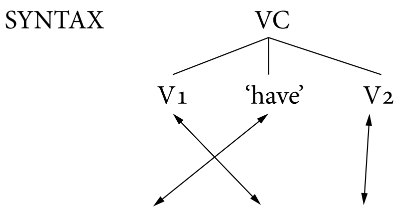
(26) . . . daß er das Buch hat lesen können/*gekonnt
that he the book has read.INF can.INF can.PST.PRT
'. . . that he was able to read the book.'

The order 2–1–3 is found in Zürich German, as shown in (27).

(27) a. dass i en gsee$_2$ ha$_1$ schaffe$_3$
that I him seen have.1S work.INF
'that I saw him work'
(M. Salzmann, p.c.)

b. Wo s aagfange$_2$ hat$_1$ rägne$_3$, . . . [ZüGe.]
when it begin.INF has rain.INF
'It began to rain.'
(Lötscher 1978)

Notice that in the case of (27a), at least, 2–1–3 puts the verb closer to its overt argument, satisfying the dependency bias. A similar case can be made for (27b), assuming that *s* 'it' is a thematic argument of *aagfange* 'begin.INF'. 2–1–3 apparently occurs only with verbs such as *anfangen* 'begin' and *sehen* 'see', which again supports the analysis of verb clusters as individuated constructions. That is, it is not correct to say that Zürich German allows the order 2–1–3 in three-verb clusters. Rather, the 2–1–3 order occurs with particular verbs, an idiosyncrasy that is straightforwardly expressed by the construction in (28).

(28)


CS $\mathbf{PAST}[\varphi_{AUX}](\mathbf{V1}'[\varphi_{V1}](\mathbf{V2}'[\varphi_{V2}])) \Rightarrow \mathbf{PAST}(\mathbf{V1}'(\mathbf{V2}'))[\varphi_{AUX}\cup\varphi_{V1}\cup\varphi_{V2}]$

where V1={'see', 'begin',...}

2–3–1 clusters with a range of verbs are well-documented in Afrikaans (Biberauer n.d.); e.g.,

(29) a. ... dat hy die medisyne kon drink het [modal]
that he the medicine could.INF drink.INF have
'... that he could drink the medicine'

b. ... dat hy hom die medisyne maak/ laat drink het [causative]
that he him the medicine make.INF let.INF drink.INF have
'... that he made/let him drink the medicine'

c. ... dat hy haar hoor roep het [perception]
that he her hear.INF call.INF have
'... that he heard her call'

d. ... dat ek haar die bokse help dra het [benefactive]
that I her the boxes help.INF carry.INF have
'... that I helped her carry the boxes'

e. ... dat die mense bly staan het [durative]
that the people remain.INF stand.INF have
'... that the people remained standing'

f. ... dat dit ophou reën het [inchoative]
that it stop.INF rain.INF have
'... that it has stopped raining'

g. ... dat hy probeer voorgee het [control]
that he try.INF pretend.INF have
'... that he tried to pretend'

h. ... dat hy die boek gaan lees het [motion]
that he the book go.INF buy.INF have
'... that he went to buy the book'

i. ... dat hy die boek loop (en) koop het [linking]
that he the book walk.INF and buy.INF have
'... that he went and bought the book'

j. ... dat hy die boek sit en lees het [linking]
that he the book sit and read have
'... that he was sitting and reading the book'

According to Biberauer (n.d.), this 2–3–1 pattern alternates with 1–2–3, and is possible only when the finite verb is a form of *hebben* 'have'. The idiosyncrasy of this pattern is captured naturally as a construction that is similar to but different in detail from (29) in Afrikaans the 2–3–1 pattern has become a distinct construction, one that is restricted to a particular finite V_1 but is very free with regards to the V_2 in the cluster. Finally, 3–1–2 is the standard order in Swiss German dialects that are uniformly 2–1, a point that is discussed further in the next section.

These data suggest that the verb clusters are constructional, and not derived by movement and adjunction. Constructions are designed to accommodate any degree of idiosyncrasy in a uniform descriptive framework. While it is possible in principle to express variation in derivational terms, using parameters, the degree of idiosyncrasy that is found in verb clusters (and I have presented only a portion of the data) would require parameters of such lexical specificity that the theory of parameters would be rendered vacuous. The parameters would effectively be a way to encode constructions in the grammar without calling them constructions. (For an extended argument along these lines, see Culicover 1999.)

Furthermore, there is evidence that a lexically restricted construction may generalize as it spreads geographically. Generalization is seen in Seiler's (2004) discussion of Swiss German verb clusters. These clusters display a number of intriguing patterns and correlations. For example, V_2-V_1 (using Standard German forms: *lesen*$_2$ *lassen*$_1$ 'make read') is not found unless V_2-MOD_1 (*lesen*$_2$ *können*$_1$ 'can read') is found, and V_1-MOD_2 is not found unless V_2-AUX_1 (*gelesen*$_2$ *haben*$_1$ 'have read') is found. At the same time, 3–1–2 (*lesen*$_3$ *kann*$_1$ *haben*$_2$ 'read can have') occurs only if the order 2–1 is fully general (across AUX, MOD and main V), suggesting a generalization of 1–2 > 2–1 to the three-verb case ('place main verb first in the sequence'). In this regard, Seiler writes (p. 13): "First, I have shown that the ordering of elements in western dialects is strictly ascending (1–2–3), but as the more we move eastwards the more the tendency for ascending ordering weakens. Second, the ordering of elements is sensitive to the category of the head. Auxiliaries tend most to be set at the right edge of the cluster. **This tendency is much weaker with modal verbs and almost absent with lexical verbs as heads of a cluster.**" (Emphasis mine—PWC)

This type of lexical idiosyncrasy is quite extensive in verb clusters. One way to interpret the Swiss data is that the 1–2(–3) pattern was originally widespread. Then the 2–1 order began as an innovation stimulated by the dependency bias, where 1 is AUX; the construction was subsequently generalized to MOD and finally V. On this view, the order 3–1–2 is a final generalization of 2–1, extending the change motivated by the dependency bias.[14]

This sensitivity to category and even to individual items is what we would expect if individual verb clusters are distinct constructions. The apparent generalization from one category to another (e.g. the generalization from auxiliaries to modals to lexical verbs) is also fully consistent with an account in which there is pressure to reduce formal complexity as marked by lexical idiosyncrasy, as discussed in §8.2.1.

A pattern similar to that found in Swiss German is seen in Heerlen Dutch (Cornips 2009). In this variety, 2–1 is preferred for V1=AUX and 1–2 is strongly preferred for

[14] Whether such changes actually occurred must unfortunately remain speculative for now. The survey of the Swiss German dialects from which Seiler's analysis is drawn dates from 2000. As far as I know there are no comparable surveys from earlier times that would provide the suitable diachronic perspective.

V1=MOD. One interpretation is that there is a change going on from 1–2 to 2–1 which is applying first to V1=AUX. This interpretation is consistent with that of Cornips (p. 212), who argues for basic 1–2 on the basis of the correlations between order preferences of individual speakers.

In West Flemish, 2–1 is required for V1=AUX and 1–2 is required for V1=MOD (Haegeman 1994). Again, this pattern suggests a shift from 1–2 to 2–1 beginning (and ending) with AUX-V.

If a variety permits 1–2 or 2–1 for a particular pair of verbs, then by iteration, it should also permit 1–(2–3) and (3–2)–1 respectively for longer combinations of these verbs. Hence we would not expect to find 3–2–1 in the absence of 2–1, or 1–2–3 in the absence of 1–2. If a variety mixes 1–2 and 2–1, as West Flemish and Afrikaans do for V1=AUX and MOD, we expect 2–3–1 where V1=AUX, V2=MOD, since V follows MOD and AUX is maximally final in the sequence. Wurmbrand (2005) identifies West Flemish as 2–3–1 in the IPP construction;[15] e.g. (30) from Haegeman (1994).

(30) ..da Valère nie nor us will$_2$-en kom$_3$-en eet$_1$
that Valery not to house want-INF come-INF has
'..that Valery did not want to come home.'

I have already suggested that 3–1–2 may be motivated by the dependency bias applying in a 1–2–3 variety. The coexistence of MOD$_1$-MOD$_2$-V$_3$ and V$_3$-MOD$_2$-AUX$_1$ in West Flemish raises the possibility that the scope bias for two modals is stronger than the dependency bias, but that the dependency bias wins when there is only one modal.

In summary, if a variety permits 1–2 or 2–1 for a particular pair of verbs, then we would expect that it permits 1–(2–3) and (3–2)-1 respectively as constructional generalizations. Crucially, we do not expect to find 3–2–1 in the absence of 2–1, or 1–2–3 in the absence of 1–2 (and we don't, as far as I can tell).

I have proposed in this section that verb clusters in which the order of elements corresponds to their scope in the semantic interpretation are the consequence of a processing bias that favors such a correspondence, while verb clusters in which the order has the elements with wider scope following are the consequence of a

[15] The Infinitivus Pro Participio construction is characterized by the fact that V$_1$ is an AUX, typically a form of 'have', and V$_2$ is MOD in the infinitival form, not the participial form. In Standard German, the IPP cluster is 1–3–2 and not the otherwise typical 3–2–1. E.g.,

(i) ...daß Peter das Buch hat lesen können
...that Peter the book has read.INF can.INF

and not

(ii) a. *daß Peter das Buch lesen gekönnt hat
that Peter the book read.INF can.PAST.PRT has
b. *daß Peter das Buch lesen können hat
that Peter the book read.INF can.inf has (Schmid 2005)

processing bias that favors positioning heads close to their dependents. The 1–2–3 clusters uniformly reflect the scope bias, while 3–2–1 clusters uniformly reflect the dependency bias. Clusters with mixed orders, such as 2–1–3 and 2–3–1, may be analyzed as reflecting both biases. Such mixed orders appear to be lexically restricted, a fact that has a natural characterization in constructional terms. Finally, it appears that lexical restrictions on a verb cluster may be relaxed or lost completely as the construction is generalized under the pressure to reduce formal complexity.

8.6 A computational simulation

Consider now the fact that in spite of differences in complexity, many different verb clusters are attested. Some clusters are far from optimal with respect to one or another bias. A natural question to ask is why complexity is maintained in the face of pressures to reduce it? Using a computational simulation, I show how a bias can lead to the reduction in the frequency of use of a construction in a population and even eliminate a construction entirely. I also show, using this simulation, that it is possible for complexity to be maintained if the social network has a suitable configuration. I hasten to note at the outset that a simulation is not a proof, but simply a potentially useful way of visualizing how an observed state of affairs may have come about.

The simulation involves a set of agents arranged in an array. These agents influence one another through contact.[16] Since we are simulating language, the agents are understood to be speakers. In the illustrations I give here, the array is 50*50. That is, there are 2500 speakers in this social network. I assume an initial state of affairs prior to contact as shown in Figure 8.5. There are three binary features which together define eight possible languages in this population of speakers. The three feature values are initially distributed randomly over the population, as shown in upper and lower right quadrants, and the lower left quadrant in Figure 8.5. One value of a feature is represented by a black square, the other by a white square. Each agent has one of the two feature values for each feature. The composite of the feature values for the three binary features is shown in the upper left quadrant. Each agent in this quadrant has one of eight possible feature combinations, represented by gradations on the gray scale from white to black. The simulation assumes that at the outset there is an even distribution of the possible feature combinations across the population, so that all eight languages are attested and have roughly the same share of the population.

At each step of the simulation, agents influence one another with respect to each feature of variation. The value of a feature of one agent will change if more of its

[16] This illustration is adapted from chapter 6 of Culicover (2013), from which Figs. 5–13 are taken. The computational model is due originally to Nowak and Latané (1994) and is described at some length in Culicover and Nowak 2003.

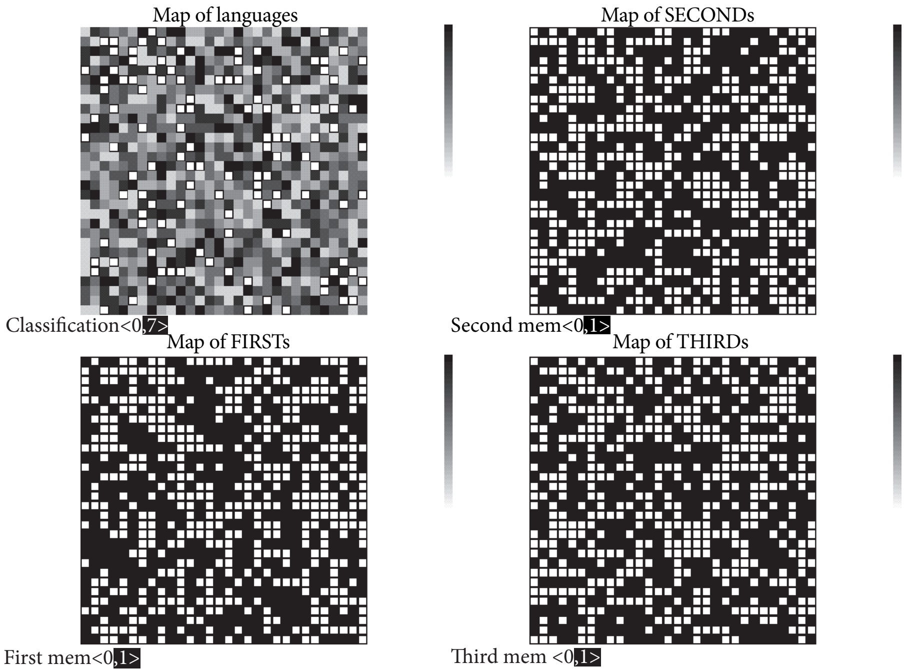


FIGURE 8.5 Feature values at initial state.

interaction partners have the other value. In this very simple simulation, all of the agents have equal influence on the others with whom they are in contact, and all of the feature values and combinations of feature values are of equal strength.[17] So, for example, if agent A is interacting with two agents, and one has the value Black and one has White, then the influence of each on A is equal. For convenience, we set the parameters of the simulation model so that A does not change from its current value unless confronted with a greater number of the alternative value. If A has White and interacts with two agents, one of which has Black and the other of which has White, A stays with White. But if the two agents both have Black, then A changes to Black. Figure 8.6 illustrates.

Figure 8.7 illustrates the distribution of languages and features after 69 steps of a simulation in which the agents in Figure 8.7 influence one another and change as a consequence of their interactions.

[17] It is possible in more complex simulations to weight the effect of distance between interacting agents, the number of agents that interact with one another, the strength of individual agents, individual resistance to change, and other parameters of the simulation model.

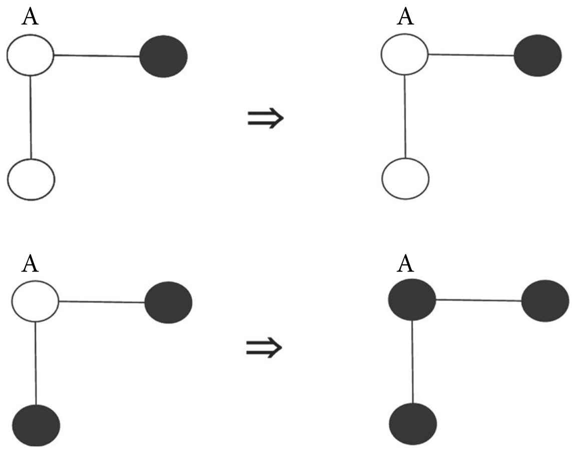


FIGURE 8.6 Interactions of neighboring agents in a network. The state of the network prior to interaction is shown to the left of the arrow, and the state of the network after interaction is shown to the right of the arrow. All interactions are computed and executed simultaneously, so it is possible that an agent may influence one agent but itself be influenced to change by other agents.

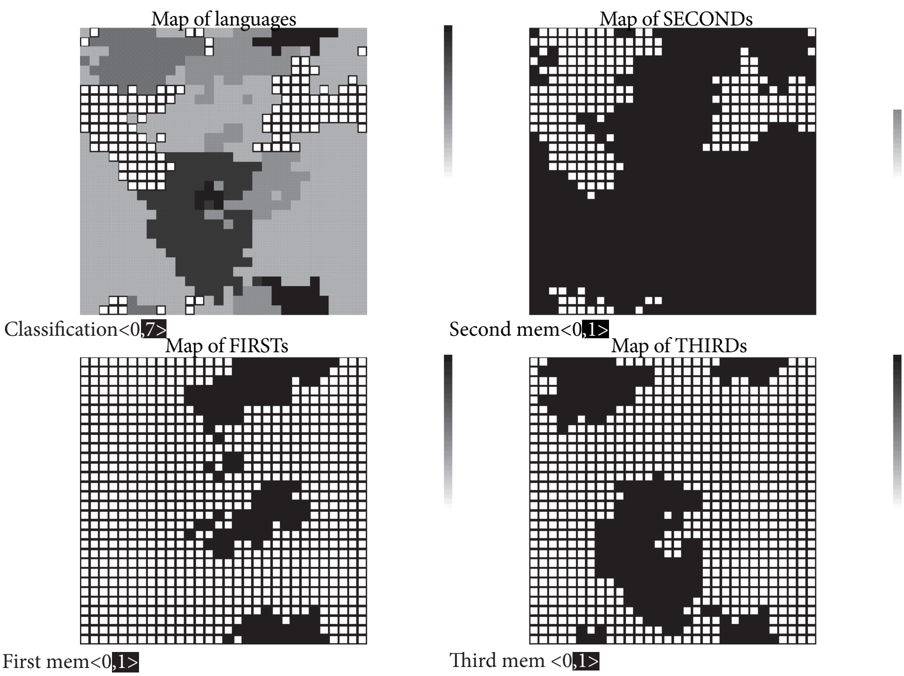


FIGURE 8.7 Feature values after step 69, showing clustering.

The display in Figure 8.7 shows the state of each agent for each feature at this point in the simulation. As a consequence of the changes from Figure 8.5, some combinations of feature values increase their distribution over the population, while others decrease. The histogram in Figure 8.8 shows the population levels of the eight languages at this point.

Notice that two languages are dead by this point, i.e. there are no agents with the relevant combination of feature values.

Now suppose that we introduce a complexity bias, such as the scope bias or the dependency bias of §8.5. Then the distribution of properties (in the present case, constructions) in the network may change in a different way than illustrated in Figure 8.7. Crucially, bias introduces an imbalance in the strength of feature values. Suppose that there is a bias in favor of Black. Now, if agent A has White, and interacts with one White and one Black, the bias in favor of Black means that Black now outranks White in the competition. So A will change to Black. Figure 8.9 illustrates, with the bias in favor of Black symbolized as a larger circle.

If the distribution of Black and White in the population is even, as in the present simulation, and the bias in favor of Black is strong enough, there is a good chance that in the end, every agent will switch to Black. How long this convergence on Black takes depends on the strength of the bias and the geography of the network of agents.

Turning back to the question of language variation, I assume that both formal and processing complexity can result in bias. In the case of formal complexity, what is in competition in the social network are the more or less general variants of a construction that express the same meaning, i.e. those with and without arbitrary conditions and exceptions. In the case of processing complexity, what is in competition in the social network are constructions that require greater or fewer computational

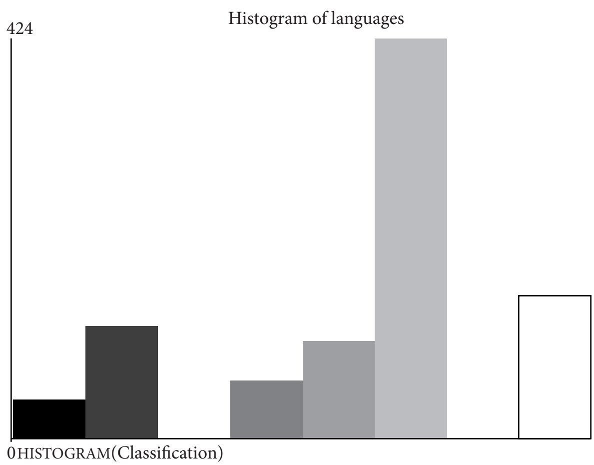


FIGURE 8.8 Number of languages of each type.

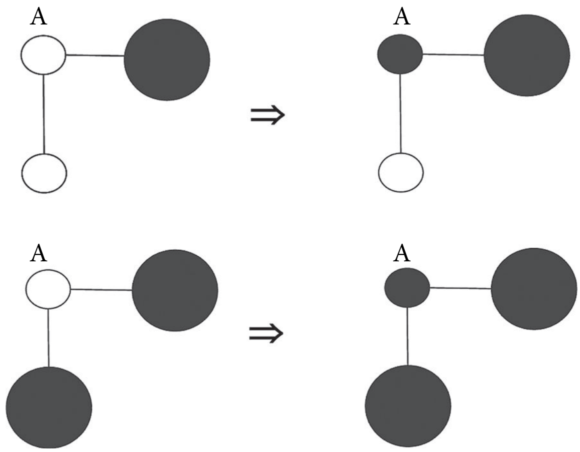


FIGURE 8.9 Interactions of agents in a network, bias on Black. The state of the network prior to interaction is shown to the left of the arrow, and the state of the network after interaction is shown to the right of the arrow. All interactions are computed and executed simultaneously, so it is possible that an agent may influence one agent but itself be influenced to change by other agents.

resources to construct a given meaning. On the view outlined here, the more general will win out over the less general, other things being equal. And the less complex in processing terms will win out over the more complex, other things being equal.

But other things are not always equal. A cluster of agents may resist external pressure to change, even when the property that is being maintained has a bias against it. The resistance to change may impede the loss of the dispreferred option. I illustrate this point with another computational simulation. I assigned a small negative bias (−1%) to one value (FIRST=Black). The initial state is that typified by Figure 8.10. Note that FIRST=Black occupies roughly the right half of the area, and THIRD=Black occupies the left half of the area. A picture of the state of the simulation after 453 steps is given in Figure 8.11.

At this point each construction has been acquired by a few speakers of the other construction.

After 2200 steps the situation is still stable, in the sense that the area occupied by FIRST=Black (the right half of the area) is still fairly large. However it is smaller than it was before, as shown in Figure 8.12, due to the negative bias. At the same time, the distribution of the THIRD property is essentially unchanged from what it was at the start of the simulation, since there is no bias affecting the distribution of this property.

It is only around steps 3400 to 3800 that the bias ultimately causes the demise of FIRST=Black in the population (Figure 8.13).

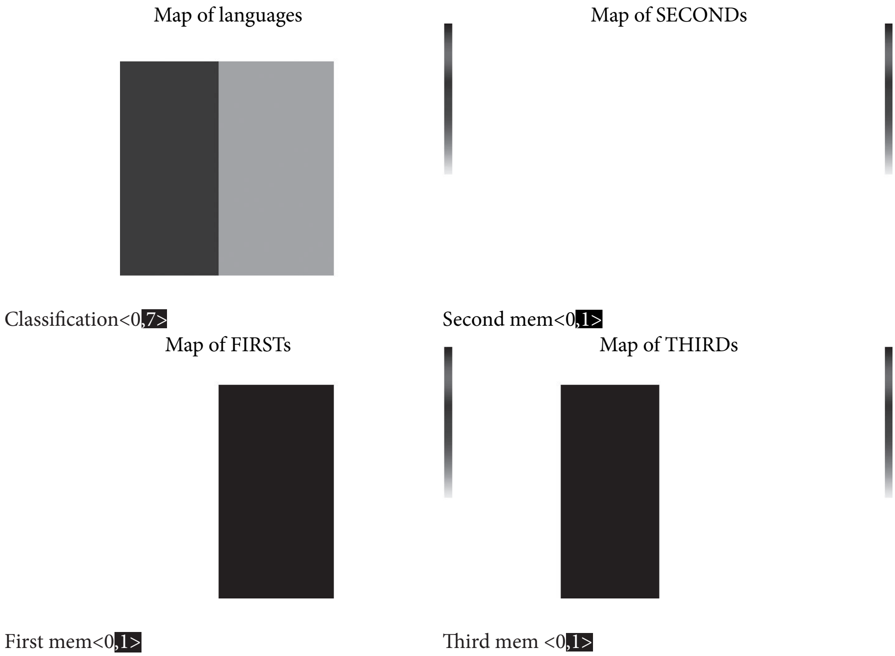


Figure 8.10 Two constructions in different regions, initial state.

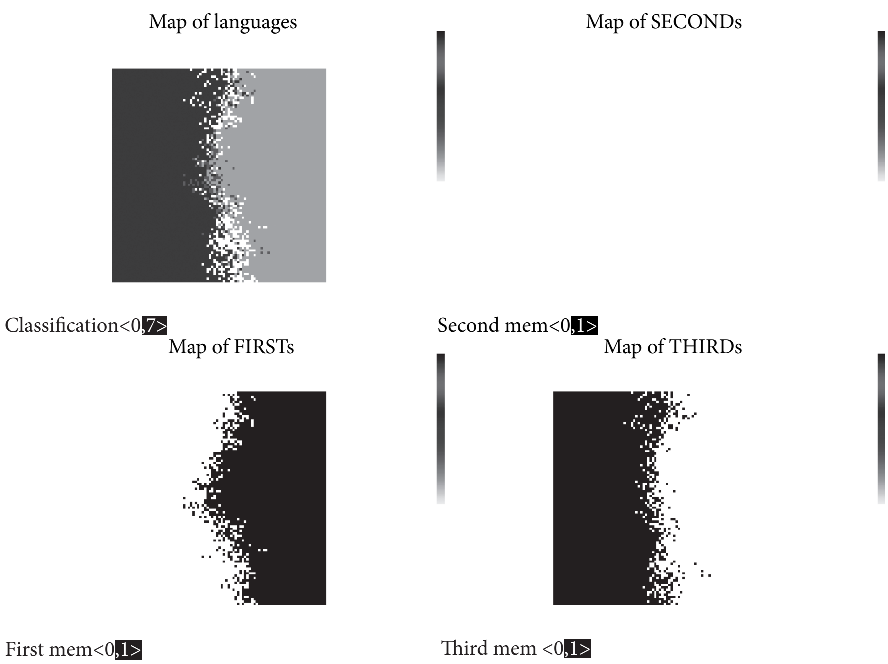


Figure 8.11 −1% bias of FIRST=BLACK. Step 453.

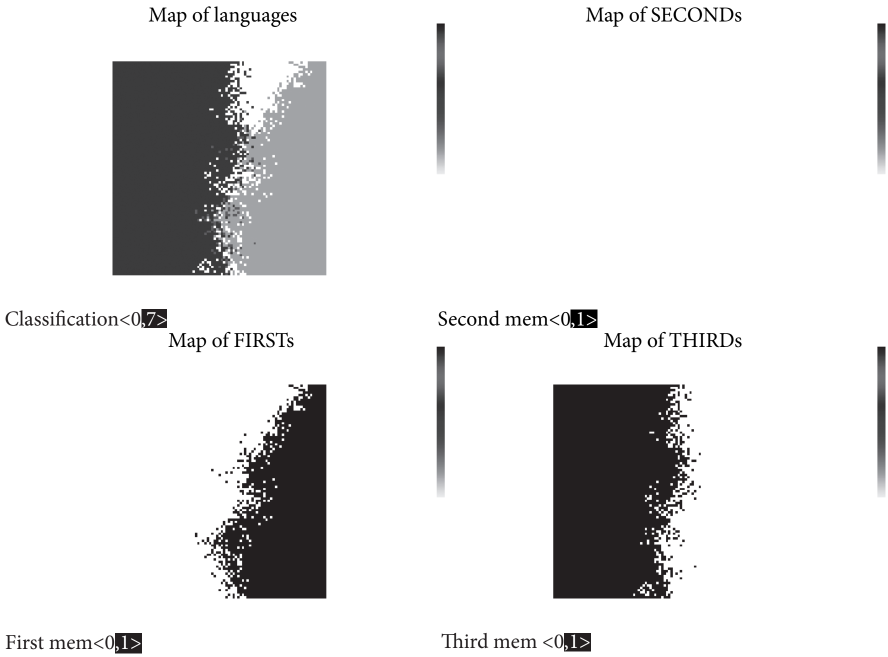


FIGURE 8.12 −1% bias of FIRST=BLACK. Step 2209.

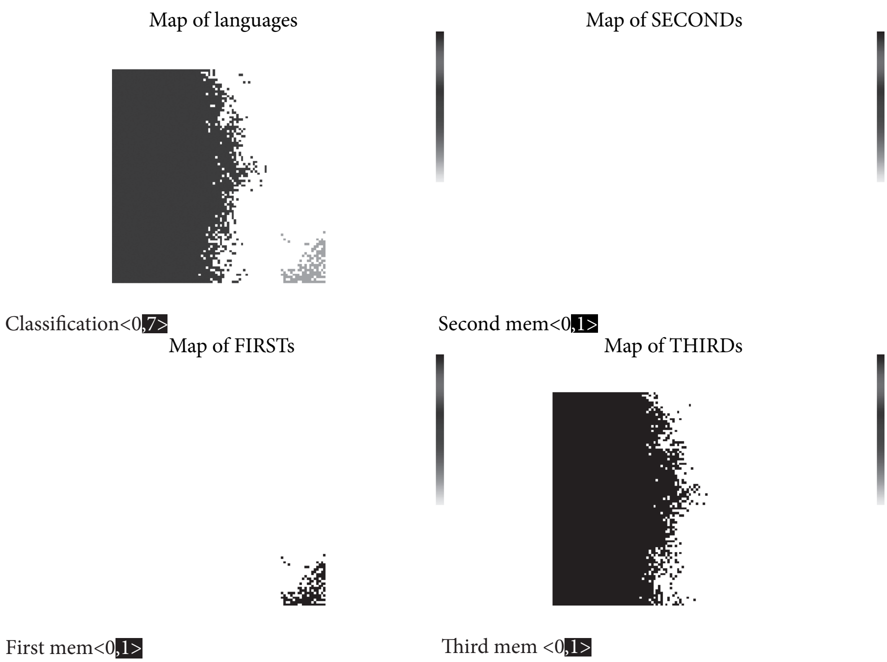


FIGURE 8.13 −1% bias of FIRST=BLACK. Step 3751.

This simulation illustrates that it is possible for a less preferred option, in this case a particular construction, to remain in the population for a substantial amount of time, and in principle forever, as long as there are conditions that continue to support it. One such condition is lack of contact of agents who have the construction with agents who have competitor constructions; another is a strong and compact cluster of agents that share the same constructions.

Thus a tight knit or relatively isolated cluster of agents may resist the effects of a bias. An agent on the edge of the cluster may change its property value as a consequence of influence of agents outside the cluster. But there is strong support from inside the cluster for that agent to change back to the predominant value in the cluster. Moreover, the agents in the center of the cluster are more resistant to influence from outside, simply because they are more insulated from external influence—they interact with fewer agents who live outside the cluster. So the influence of the bias is less, to the extent that there are well populated clusters that can resist it. The bias adds strength to a particular value, and therefore may induce an agent with another value to switch to it. But what also contributes to the strength of a construction is the number of agents that have it, compared with the number that have the other value.

Finally, what is not reflected in the computational simulation but almost certainly has a significant effect in real social networks is the frequency of use of a dispreferred alternative.[18] This last factor is a plausible explanation for the persistence of irregular verbal morphology, for example, in languages such as English or German.

The computational simulation in this section illustrates a number of points. First, other things being equal, a complexity bias can affect the distribution of properties represented in the social network. Given two constructions of different complexity, the more complex construction is likely to be less represented in the population than the less complex construction. Second, other things are not always equal, and the configuration of the network affects how agents interact with one another. This in turn affects how properties will be propagated in the network. Third, a weak complexity bias may be resisted by clusters of agents, so that a more complex construction may coexist in the network with another less complex construction, perhaps indefinitely.

8.7 Summary and conclusions

In this chapter I have suggested a way of accounting for word order and word order variation in terms of complexity. I have argued that it is useful to consider word order variation in terms of competing constructions. Other things being equal, the less complex construction is preferred to the more complex alternative. When there are

[18] For an extensive examination of the role of frequency in explaining regularity and irregularity in language, see Bybee (2007).

multiple factors contributing to complexity, one construction may predominate in one language variety and a competing construction in another language variety as a consequence of social factors.

I discussed two types of complexity: formal complexity, which has to do with how regular a language is, and processing complexity, which has to do with the computational resources required to relate syntactic configurations and their interpretations. Both types of complexity play a role in language change and variation. However, although it is commonplace to assume that change occurs in order to reduce complexity, I showed that using this notion to understand actual cases of change is far from straightforward. Using the example of English infinitival relatives, I argued that in the case of formal complexity, a change in the grammar that increases regularity in one respect may introduce idiosyncrasy in another respect.

I then turned to Continental West Germanic verb clusters to illustrate how processing complexity may contribute to word order variation. In this case, I argued that it is useful to consider each cluster with a particular order of verbs as a phrasal construction, that is, a form–meaning correspondence. By assumption, all orderings of the daughters of a phrase are available to speakers, although they will actually produce and fully accept those for which there is substantial positive evidence.

Furthermore, I showed that different orderings of the verbs in a cluster differ in complexity with respect to different measures. Those clusters in which the operators, that is, the elements with wide scope, precede their arguments are less complex with respect to the scope bias. Those clusters in which the arguments precede the operators are less complex with respect to the dependency bias, since in these cases, the lexical heads are closer to their dependents.

Given that there are multiple dimensions of complexity, there is no single verb cluster that is least complex. Rather, there are simultaneous pressures to maintain a particular ordering and to change to another order. The availability of all orderings allows for this latter possibility, even where the alternative order is not attested in the language. In a computational simulation, I illustrated how the biases might operate within a social network to create subgroups of speakers who favor one or another alternative constructional variant.

Of course, the situation in the real world is much more complex than is conveyed by the particular cases discussed here. For one thing, there are many factors that enter into word variation, not just the scope bias and the dependency bias. For example, Wasow (2002) shows that word order variation in the English VP is sensitive to grammatical weight, discourse newness, and collocation frequency, among other factors. It is possible that some if not all of these factors are also responses to complexity, on dimensions other than the scope bias and the dependency bias.

Finally, I argued that it is useful to characterize each of the possible verb clusters as an individual construction. In general, it is not possible to say that there is a particular ordering of all of the two- and three-verb clusters of a particular variety. Rather, the evidence shows that actually occurring verb clusters must in some cases be defined in terms of specific verbs or verb classes. In the case of Swiss German verb clusters, for example, particular orders (e.g. 2–1) occur in one dialect only when the second verb is AUX. The Swiss German case is particularly interesting because it can be seen that in neighboring dialects, the 2–1 order is extended to clusters in which the second verb is MOD, and in some dialects, to all two-verb clusters. Both the lexical idiosyncrasy of particular clusters and the generalization of particular clusters to more inclusive lexical classes has a natural characterization in terms of constructions.

9

Complexity trade-offs: a case study

KAIUS SINNEMÄKI

9.1 Introduction

One widely accepted hypothesis among linguists is that complexity in one part of the grammar of a language correlates with simplicity in another (henceforth 'the trade-off hypothesis'), allowing all languages to manifest roughly the same degree of complexity (e.g. Hockett 1958: 180–181; Edwards 1994: 90–91; Bickerton 1995: 35, 67; Aboh and Smith 2009: 4).[1] However, recent typological studies have suggested that only some linguistic variables participate in complexity trade-offs (e.g. Fenk-Oczlon and Fenk 1999; Shosted 2006; Sinnemäki 2008; Miestamo 2009; Parker 2009). Based on these studies, it seems that the trade-off hypothesis is not an all-encompassing principle in human languages. What makes the hypothesis still a matter of interest is that it raises three interesting questions: first, To what extent do complexity trade-offs occur?; second, What kind of complexity might be involved?; and third, What causes these purported effects?

My purpose here is to study the trade-off hypothesis in one narrow grammatical domain, namely, the marking of the basic participants in the sentence, or, core argument marking (for more discussion of the hypothesis, see Moran and Blasi, this volume, chapter 11). Following the discussion in Comrie (2011), core argument marking discriminates between the more agent-like (A) and the more patient-like (P) core arguments of a two-place transitive predicate. This domain provides a feasible *tertium comparationis* for the purpose because the coding devices used in this domain—case marking, agreement, and rigid word order—tread the same functional area and, thus, they have often been expected to correlate negatively (e.g. Kiparsky 1997; Blake 2001; Malchukov and Spencer 2009). I focus here on case

[1] I would like to thank the editors, Frederick Newmeyer and Laurel Preston, for several valuable comments that helped to improve the chapter. I am also grateful to the University of Tampere and the Helsinki Collegium for Advanced Studies, University of Helsinki, for making this research possible.

marking and rigid order, since agreement does not seem to correlate with the other two coding devices in a meaningful way, as I argue in Sinnemäki (2008). My research questions are: Does complexity in case marking correlate with simplicity in rigid order cross-linguistically (and vice versa)? and What kind of complexity might be involved?

I perform three experiments to approach these questions. In two experiments I approach case marking and rigid order as 'resources' that are available in languages to build utterances; in the third I focus on the ways that the use of these coding devices may be 'regulated' via dependencies between other grammatical systems in languages, adapting the approach of Aikhenvald and Dixon (1998). 'Resources' and 'regulations' are notions that Dahl (2004: 40–41, 2008: 154) proposes in connection to applying the notion of complexity to language: the former deals with the inventory of linguistic units and constructions (e.g. morphemes and phrases) that are used in a language to build utterances, while the latter deals with 'the constraints and requirements that the system puts on its users', similar to 'rules of grammar' but more general in scope and less loaded in meaning (Dahl 2008: 154). Data from a stratified sample of 50 languages provide evidence for a strong negative correlation between case marking and rigid order in terms of complexity, but only in terms of resources, not regulations. I assume that the attested trade-off is best explained by the principle of economy (see §9.4).

The chapter is organized as follows. In section 9.2, I develop an approach for measuring grammatical complexity in different languages and explain how I measure the complexity of case marking and rigid order. In section 9.3, I define the object of the case study, discuss the sample, and present the results of the experiments. Sections 9.4 and 9.5 are devoted to a discussion of the results and a brief conclusion respectively.

9.2 A cross-linguistic approach to complexity

In this section I propose a way of approaching grammatical complexity from a typological perspective. I take as a starting point that complexity could be characterized at a general level in terms of the number and variety of parts and their interrelations in a system—there seems to be a rather wide agreement about this idea (e.g. Simon 1996: 183–184; Rescher 1998: 1; Hübler 2007: 10). In the following subsections I explain how this general characterization could be made more precise for measurements.

9.2.1 *Measuring and delimiting complexity in cross-linguistic research*

In this section I discuss two ways of delimiting complexity for cross-linguistic research (§9.2.1.1) and explain how description length could be used as a basis for complexity measures in typology (§9.2.1.2).

9.2.1.1 *Delimiting complexity* A crucial question in language complexity research is whether to take a usage-based or a purely grammatical description-based approach. Given the former, the complexity of different structures, such as case marking patterns, would be based on their on-line difficulty in language use (Kusters 2003). Given the latter, the formulation of complexity would be based on the number and variety of the parts of the grammatical description and their interactions (Dahl 2004; Miestamo 2008).

Following Dahl (2004), Miestamo (2008, 2009), and Sinnemäki (2011), I consider the latter approach superior as a basis for a general typological approach to complexity. The most important reason for this assumption is that different user-based criteria do not necessarily lead to the same complexity measurement: the speaker, the hearer, the first language acquirer, and the second language learner do not all find the same things easy or difficult (see Miestamo 2008; Sinnemäki 2011 for details).

Another important delimitation of complexity is that any complexity measurement needs to focus on local, rather than global, complexity, as suggested by Miestamo (2008), (though see Ross, this volume, chapter 10 for a different point of view). The former concerns the complexity of some subpart of a grammar, for instance, the tense-aspect system, while the latter concerns the overall complexity of a language's full grammar.

There are a number of reasons for concentrating on local complexity. I mention here just two, which Miestamo (2008) calls the problems of 'representativity' and 'comparability' (see Deutscher 2009; Sinnemäki 2011 for further discussion). First, it is impossible at the present time to provide a comprehensive description of a grammar of even one language. How then could we compare two languages in terms of their relative overall complexity? Second, it is unclear what the relative contribution of different subparts of grammar to global complexity would be. For example, how should we weigh the complexity of a tense-aspect system in relation to the complexity of rigid order and how should we estimate their contribution to global complexity? Fortunately, since my focus is on complexity trade-offs in a particular grammatical domain, I necessarily concentrate on local complexity and avoid these problems related to global complexity.

9.2.1.2 *Description length* I propose here that complexity be measured in terms of notions provided by information theory (e.g. Shannon 1948). Information-based measures of complexity have been popular in different disciplines and recently also in linguistics (e.g. Bane 2008; Juola 2008). In algorithmic information theory a common metric of complexity is what is called 'Kolmogorov complexity', namely, the length of the shortest possible description required to specify an object (e.g. Kolmogorov 1965; Chaitin 1987).

A characteristic of Kolmogorov complexity is that the more disorderly a string, the longer its description, and hence the greater its complexity. However, it is not maximum disorder that is of interest when measuring grammatical complexity, but rather the set of regularities within the system, that is, its structure (see Dahl 2004: 24 and Miestamo 2009: 81 for similar arguments). Descriptive length of the set of regularities is what Gell-Mann (1995) calls 'effective complexity'. In the rest of the chapter, I adopt 'effective complexity' as the basis for complexity measures, meaning that complexity is relative to the length of description required by the system's regularities.

Description length can be applied to measuring the effective complexity of linguistic patterns as well. The idea is that the more patterns a linguistic entity contains, the longer its description, and hence the greater its complexity. For example, a verb in Maybrat (Dol 1999: 69) contains maximally one inflectional morpheme (1), while a verb in Turkish contains many more, as in (2) (Göksel and Kerslake 2005: 74; see Bickel and Nichols 2011 for estimating the degree of inflectional synthesis). A partial description of the verb and its inflectional synthesis in Maybrat could be 'person. gender-ROOT' and in Turkish 'ROOT-reciprocal-causative-causative-passive-negation-possibility-imperfective-evidential.copula-person.number-generalizing.modality'. Verbal inflectional morphology is thus less complex in Maybrat than in Turkish, assuming effective complexity.

(1) Fane y-tien. (Maybrat)
pig 3m-sleep
'The boar sleeps.'

(2) Döð-üþ-tür-t-ül-me-yebil-iyor-muþ-sunuz-dur. (Turkish)
beat-recp-caus-caus-pass-neg-psb-ipfv-evid.cop-2pl-gm
'It is presumably the case that you sometimes were not made to fight.'

Note that the number of inflectional categories on the verb can be ranked on a complexity scale—the more categories, the more complexity—and therefore languages with different complexity values with respect to verbal inflection can be compared (see Shosted 2006 for details). In this way, description length can be applied to measuring effective complexity of linguistic patterns, albeit without recourse to compression algorithms that are sometimes used for measuring description length (see e.g. Juola 2008), but rather using linguists' descriptive tools as the basis (see Miestamo 2008; Sinnemäki 2011 for a more detailed discussion).

9.2.2 *Measuring the complexity of case marking and rigid order*

In this section I propose three ways of examining the complexity of case marking and rigid order, each focusing on a different way of describing grammar. The first two (§9.2.2.1 and §9.2.2.2) deal with 'resources', while the third (§9.2.2.3) deals with 'regulations' in the language system (see §9.1 for definitions), the requirements and constraints that the system imposes on its users.

It is possible to relate resources and regulations to the general characterization of complexity that I gave in the beginning of §9.2, in the following way: resources are roughly related to the number and variety of parts in a system, because they are the building blocks that make up utterances or "the set of possibilities that a system offers its users" (Dahl 2008: 154). Requirements, restrictions and constraints, on the other hand, are roughly related to the interactions between the system's parts, such as when animacy interacts with case marking to restrict overt case marking to animate referents only.

While case marking is clearly a matter of resources, rigid order straddles the borderline between both resources and regulations. On the one hand, rigid order is a matter of restricting the order of utterance elements and, therefore, a matter of regulations (see §9.3.1 for the diagnosis of rigid order). On the other hand, rigid order is a matter of resources as well: It is a property of constructions in a language (see §9.3.1 for more discussion) and therefore relevant to building the meaning of utterances, parallel to other resources in a language and, therefore, a matter of resources. For the purpose of this chapter, I consider rigid order as a resource on a par with case marking for marking the core arguments.

9.2.2.1 *Complexity as overt coding* The simplest way to describe case marking and rigid word order is to describe their availability in a language. I call this approach 'complexity as overt coding', because it is based on the mere presence or absence of the coding devices.

But how is overt coding related to effective complexity? In Sinnemäki (2011: 37–39) I assume that overt case marking is more complex than its absence, since overt case marking requires a longer minimal description than the absence of case marking (see §9.3.1 for the diagnosis of case marking). I also assume that rigid order has a greater effective complexity than non-rigid order (in core argument marking), because rigid order imposes a restriction on argument position and, therefore, requires a longer minimal description than non-rigid order.

9.2.2.2 *Complexity as variety of marking patterns* Approaching complexity as overt coding is a useful first step in measuring grammatical complexity, but it is based on a coarse-grained description of the grammar. A more fine-grained description would take into account possible variation in the marking patterns, providing a more detailed account of the resources available for argument marking.

For my second way of examining complexity, I propose to account for this variation by examining the variety of marking patterns available for case marking and rigid order in argument marking. As for case marking, this means examining the number of different case forms that are used for distinguishing the arguments from one another (excluding morphophonological variation). For instance, Trumai uses only one overt case marker, the ergative case marker *-k*, in core argument marking (Guirardello-Damien 2010), whereas Semelai uses the clitic *la=* to mark the ergative argument and the clitic *hn*= to mark the absolutive argument (Kruspe 1999: 149–152). Jaqaru, on the contrary, uses five different marking patterns in different contexts: dropping the final vowel of the noun root (when the noun occurs near the verb to mark the subject and when the noun is removed from the verb to mark the object), dropping the second vowel of the noun root if the noun is unsuffixed and preverbal to mark the object, and the accusative suffix *-ha* or the directional *-ru* 'to' to mark the object (Hardman 1966, 2000). Based on these facts, I assume that the length of description required to account for the variety in case marking is smallest in Trumai, greater in Semelai, and the greatest in Jaqaru.

As for rigid order, I examine the number of different rigid order patterns that are used for argument discrimination within a language. For instance, Pirahã uses one rigid order throughout (subject-object-verb; Everett 1986), while Welsh uses rigid verb-subject-object order in one construction and rigid auxiliary-subject-verb-object order in another (see §9.3.1 for details; Borsley *et al.* 2007). As far as the number of separate rigid orders is concerned, I assume that it requires a longer description to account for the two rigid orders in Welsh than the single rigid order in Pirahã.

9.2.2.3 *Complexity as system dependencies* My third way of examining complexity focuses on the ways in which case marking and rigid order may interact with other grammatical systems in the language, such as tense or animacy. For the purpose of examining these interactions, I adapt the approach of Aikhenvald and Dixon (1998) for describing dependencies between grammatical systems (henceforth, system dependencies).

The idea of system dependencies is that the choices available, for instance, in the case system, may depend on the choices made, say, in the number system. A good example is West Greenlandic, which distinguishes between the ergative and absolutive cases in the singular but not in the plural (see Table 9.1).

TABLE 9.1. Ergative and absolutive case forms in West Greenlandic (Fortescue 1984: 206)

	Singular	Plural
Ergative	-(u)p	-(i)t
Absolutive	-q/-t/-k/zero	-(i)t

Another typical example of a system dependency is when the choices available in the case system depend on the animacy of the referent, so that there are more case distinctions when the referent is animate than when it is inanimate (Aikhenvald and Dixon 1998: 67). For instance, the arguments are obligatorily distinguished from one another in Kannada when the object referent is animate (Lidz 2006). The situation is analogous for rigid order. In Lakhota, rigid order is used when both arguments are animate (VanValin 1977), but in West Greenlandic it is used only when both arguments are plural (Fortescue 1984: 181).

For the purpose of this experiment I apply a slightly more inclusive notion of system dependency than Aikhenvald and Dixon (1998). A dependency between grammatical systems may also occur when a choice in one system depends on a choice in another system. In Diyari the form of the ergative case depends on the gender of the referent noun (among other things), so that male personal names are marked with *-li* and female personal names with *-ndu* (Austin 1981: 48–49). The idea here is that it is not the number of choices available in the case system that depends on the gender system, but it is rather the choice of a particular case form that depends on gender.

In the experiment, I examine the different ways in which case marking and rigid order depend on other grammatical systems. In some languages, the marking pattern depends on just one grammatical system, for instance the number system, as in West Greenlandic. In others, the marking pattern depends on more grammatical systems. In Kannada, accusative case marking is obligatory for animate objects, inanimate plural objects, and for displaced objects in clause-initial focus position (Sridhar 1990; Lidz 2006). The case marking system of Kannada thus depends on three grammatical systems, namely, animacy, number, and information structure.[2]

But what is the rationale for considering a marking pattern more complex if it depends on some other grammatical system? I assume that it requires a longer description to account for a marking pattern which depends on some other grammatical system compared to a marking pattern that is not restricted in that way and in that sense, case marking in Kannada has greater effective complexity compared to, for instance, case marking in Hungarian (Kenesei *et al.* 1998), which does not depend on other grammatical systems.

9.3 A case study of core argument marking

In this section I define and delimit the object of the case study (§9.3.1), discuss sampling (§9.3.2), and present the results of the experiments (§9.3.3).

9.3.1 *Key definitions*

In this case study I approach complexity trade-offs in a typological way, that is, by classifying grammatical variables in terms of complexity, forming a representative sample of the world's languages, and assessing whether the grammatical variables correlate negatively with one another in the sample languages. In typology an

[2] Sometimes the sources state simply that case marking or rigid order is used when there is a danger of confusing who is doing what to whom (e.g. Hardman 2000 for Jaqaru). In these situations I postulate that the marking pattern depends on one grammatical system, namely the animacy of the arguments. Typically animacy of the referent helps in disambiguation, since the subject tends to be animate and definite and the object inanimate and indefinite (e.g. Comrie 1989), but otherwise animacy may be of little help. Although other factors might be relevant as well, this could not always be verified from the sources.

important requirement is to select grammatical variables that are comparable across languages. In statistical research, on the other hand, an important requirement is to select variables that have a theoretically motivated connection to each other in order to exclude chance correlations. Both of these requirements are met when variables are chosen from the same functional domain. A functional domain is defined by Givón (1981) as a domain that marks similar semantic-pragmatic functions in different languages, such as possession or core arguments. Core argument marking serves my purpose well, since it is one of the most thoroughly studied cross-linguistic properties of grammar.

For practical and methodological reasons it is generally necessary to delimit the typological data in various ways. The following three delimitations on core argument marking are adapted from Siewierska (1998), who presents a more thorough set of delimitations for her investigation of word order variation. My study resembles that of Siewierska's, but with the major difference that she focuses on word order variation of subject, object and verb in general, while I focus solely on the role of word order in argument discrimination.

My first delimitation is to study case marking and word order in simple clauses, that is, in simple active affirmative indicative main clauses with a transitive predicate. While polarity, voice, mood, or clause type can affect argument marking (e.g. Palmer 1994: 61–62; Siewierska 1998: 477–478), they are excluded to keep the workload bearable.

Second, I concentrate on fully lexical noun arguments. While pronouns are more frequent in discourse than nouns, their higher frequency makes them more conservative and thus more resistant to universal effects, such as morphological levelling (Bybee and Thompson 1997; Moravcsik 2011). English is a good example, having lost case marking of nouns but not of pronouns (*she carried the baby/her*). So if there is a tendency in languages to compensate for the lack of case marking by rigid order, it is likely that such a compensating tendency will manifest itself historically earlier in fully lexical nouns than in pronouns, given the greater conservatism of case marking of pronouns.

Third, I exclude arguments modified by optional modifiers, such as adjectives, quantifiers, or subordinate clauses, because modified arguments may behave differently from non-modified ones. For instance, in Cora, the accusative case is used only on some numerals and quantifiers, but not elsewhere (Casad 1984: 230).

Let us now turn to the definitions of 'case marking' and 'rigid word order'. I define 'case marking', following Nichols (1992), as any dependent marking of a core argument. An argument is dependent-marked when its form is distinct from that of the other argument due to overt coding. This definition is indifferent to the way overt marking is realized, encompassing marking by affixes, isolating formatives, and tonal and morphophonological alternations (see Bickel and Nichols 2007: 172–193 for more details).

Rigid word order can be identified when interchange in the position of two arguments triggers a change in the thematic interpretation of the sentence. For instance, changing the position of the arguments in the English sentence *the boy kissed the girl* to *the girl kissed the boy* changes the thematic interpretation of the whole sentence. In other words, the only way of interpreting each sentence is to interpret *the boy* as the one who kissed in the first sentence, but as the one who was kissed in the second sentence. Note further that it is the interchange of positions that is crucial here, not merely a change in the order of arguments (pace e.g. Primus 1999: 132). For instance, in the sentence *Mary John kissed* the order of arguments has changed in comparison to the sentence *John kissed Mary*, but this word order change does not change the thematic interpretation of the sentence, because the positions of the arguments have not been interchanged.

In identifying rigid word order, I exclude arguments that occur outside the clause proper, that is, those that are separated from the rest of the clause by a pause, as in Pirahã (3) (Everett 1986: 201–202) and/or marked by a resumptive pronoun in the clause, as in Middle Atlas Berber (4) (Ouali 2006: 38–39). A roughly similar delimitation of the data was used by Siewierska (1998) in her study of word order variation.

(3) a. (Pirahã)

Ti	xíbogi	ti-baí.
1	milk	drink-ints

'I drink milk.'

b.

Xíbogi	ti	ti-baí.
milk	1	drink-ints

*'I drink milk.' (possible only using a pause after *xíbogi*)

(4) a. (Middle Atlas Berber)

Yugh	Moha	aksum.
3sg.sbj.bought.pfv	Moha	meat

'Moha bought meat.'

b.

Ahslaf	yukku-t	Ali.
grass	3sg.sbj.cut.pfv-it	Ali

'The grass that Ali cut.'

Also, I consider rigid word order as a property of constructions, not of the language as a whole, so there can be several separate rigid order types within a language. Let me give two types of examples. In some languages the predicate may be discontinuous, as in Welsh. In this language, verbal categories are expressed via inflection on the main verb (5a) or by a periphrastic construction (5b), which contains a finite copula (or an auxiliary) and the main verb in a non-finite form separated from the finite verb (Borsley *et al.* 2007: 33). As far as discontinuous orders are rigid, such as in (5b), I analyze them as separate rigid order types within the language (note that Siewierska 1998 excluded the latter type of constructions from her study).

(5) a. Gwelodd Rhiannon ddraig. (Welsh)
see.pst.3sg.sbj Rhiannon dragon
'Rhiannon saw a dragon.'

b. Mae Rhiannon wedi gweld draig.
be.prs.3sg.sbj Rhiannon pfv see.inf dragon
'Rhiannon has seen a dragon.'

In other languages, interchanging the position of the arguments may trigger special morphological marking. In these situations I postulate that both the morphologically unmarked word order and the morphologically marked word order are separate rigid order types within the language. For instance, in Trumai (Guirardello-Damien 2010: 204–209), the P (the more patient-like argument) usually occurs in an immediately preverbal position (6a), but if it is moved away from this position, the clause is always, and only then, marked with a post-verbal particle *ke* (6b) (Guirardello-Damien 2010: 208). Therefore, I postulate that Trumai has two separate rigid order types, namely APV (Agent-Patient-Verb), which is not marked with the particle *ke*, and PAV (Patient-Agent-Verb), which is marked with the particle *ke*.

(6) a. Ine-k ka_in atlat-∅ mapa. (Trumai)
3-erg foc/tns pan-abs break
'He (focus) broke the pan.'

b. Atlat-∅ ka_in ine-k mapa ke.
pan-abs foc/tns 3-erg break ke
'He broke the pan (focus).'

These definitions and delimitations provide a narrow but cross-linguistically comparable domain for the case study.

9.3.2 *Sampling procedure and the sampled languages*

The data for the case study come from a stratified random sample of 50 languages.[3] All languages were first classified genealogically in genera, subfamilies, and families following the taxonomy of the *World Atlas of Language Structures* (henceforth *WALS*; see Dryer 2011 for the taxonomy). A genus is a genealogical unit roughly corresponding to the branches of Indo-European (e.g. Germanic, Romance). Languages and genera were then areally classified in six macro-areas following Dryer (1992): Africa (AFR), Eurasia (EUR), Southeast Asia-Oceania (SAO), Australia-New Guinea (ANG), North America (NAM), and South America (SAM). The sample was formed by taking randomly from each of these macro-areas a number of genera that

[3] The sample is largely based on Sinnemäki (2008, 2011: 117–121), with minor changes in the selection of languages.

TABLE 9.2. **Taxonomy and sources for each sample language**

Area	Language	Genus	Family	Source
AFR	Berber (Middle Atlas)	Berber	Afro-Asiatic	(Penchoen 1973; Ouali 2006)
	Somali	Eastern Cushitic	Afro-Asiatic	(Saeed 1999)
	Khoekhoe	Central Khoisan	Khoisan	(Hagman 1977)
	Lunda	Bantoid	Niger-Congo	(Kawasha 2003)
	Kisi	Southern Atlantic	Niger-Congo	(Childs 1995)
	Ngiti	Lendu	Nilo-Saharan	(Kutsch Lojenga 1994)
	Nubian (Dongolese)	Nubian	Nilo-Saharan	(Armbruster 1960; Dimmendaal 2010)
EUR	Kannada	Southern Dravidian	Dravidian	(Sridhar 1990; Lidz 2006)
	Welsh	Celtic	Indo-European	(Borsley *et al.* 2007)
	Georgian	Kartvelian	Kartvelian	(Harris 1981)
	Hungarian	Ugric	Uralic	(Kenesei *et al.* 1998; Rounds 2001)
SAO	Semelai	Aslian	Austro-Asiatic	(Kruspe 1999)
	Malagasy	Barito	Austronesian	(Rasolofo 2006)
	Nias	Sumatra	Austronesian	(Brown 2001)
	Hmong Daw	Hmong-Mien	Hmong-Mien	(Fuller 1985; Cresswell and Snyder 2000)
	Qiang	Qiangic	Sino-Tibetan	(LaPolla and Huang 2003)
	Thai	Kam-Tai	Tai-Kadai	(Diller 1993; Iwasaki and Ingkaphirom 2005)
ANG	Gooniyandi	Bunuban	Australian	(McGregor 1989, 1990)
	Alawa	Maran	Australian	(Sharpe 1972)
	Diyari	Pama-Nyungan	Australian	(Austin 1981)
	Daga	Dagan	Dagan	(Murane 1974)
	Kuot	Kuot	Kuot	(Lindström 2002)
	Yimas	Lower Sepik	Lower Sepik-Ramu	(Foley 1991)
	Oksapmin	Oksapmin	Oksapmin	(Loughnane 2009)
	Skou	Western Skou	Skou	(Donohue 2004)

(*continued*)

TABLE 9.2. **Continued**

	Lavukaleve	Lavukaleve	Solomons East Pap.	(Terrill 2003)
	Klon	Greater Alor	Timor-Alor-Pantar	(Baird 2008)
	Arapesh	Kombio-Arapesh	Torricelli	(Conrad and Wogiga 1991)
	Korowai	Awju-Dumut	Trans-New Guinea	(van Enk and de Vries 1997)
	Maybrat	North-Central Bird's Head	West Papuan	(Dol 1999)
	Yelî Dnye	Yele	Yele	(Henderson 1995; Levinson 2006)
NAM	Cree (Plains)	Algonquian	Algic	(Dahlstrom 1991)
	Greenlandic (West)	Eskimo	Eskimo-Aleut	(Fortescue 1984)
	Tzutujil	Mayan	Mayan	(Dayley 1985; Duncan 2003)
	Choctaw	Muskogean	Muskogean	(Broadwell 2006)
	Slave	Athapaskan	Na-Dene	(Rice 1989)
	Miwok (Southern Sierra)	Miwok	Penutian	(Broadbent 1964; Davis 2002)
	Lakhota	Siouan	Siouan	(Van Valin 1977)
	Cora	Corachol	Uto-Aztecan	(Casad 1984)
	Nuuchahnulth	Southern Wakashan	Wakashan	(Waldie 2001; Wojdak 2005)
SAM	Jaqaru	Aymaran	Aymaran	(Hardman 1966, 2000)
	Hixkaryana	Cariban	Cariban	(Derbyshire 1979, 1985)
	Ika	Aruak	Chibchan	(Frank 1990)
	Pirahã	Mura	Mura	(Everett 1986)
	Shipibo-Konibo	Panoan	Panoan	(Valenzuela 1997, 2003)
	Yagua	Peba-Yaguan	Peba-Yaguan	(Payne 1985; Payne and Payne 1990)
	Quechua (Imbabura)	Quechuan	Quechuan	(Cole 1985)
	Trumai	Trumai	Trumai	(Guirardello 1999; Guirardello-Damien 2010)
	Urubú-Kaapor	Tupi-Guaraní	Tupian	(Kakumasu 1986)
	Berbice Dutch Creole	Creoles	Creoles	(Kouwenberg 1994)

is in proportion to the total number of genera in each macro-area, following Miestamo (2005: 31–39).[4] All macro-areas are thus represented in the same proportion of genealogical variation at the genus-level. Table 9.2 provides information about the taxonomy and the primary sources for each language and Figure 9.1 shows the languages on a world map.

The sample was genealogically stratified so that no two languages came from the same genus, subfamily, or family. I allowed minor concessions at the family-level in some macro-areas due to unavailability of data at the time of sampling. One creole was also sampled by virtue of the centrality of creoles in the recent complexity debates (e.g. McWhorter 2001; DeGraff 2001; Aboh and Smith 2009). Given the above, I feel that the sample of the world's languages is reasonably representative. Finally, data for each language were drawn from grammar descriptions and works dealing with the issues at hand, as well as from personal communication with language experts.

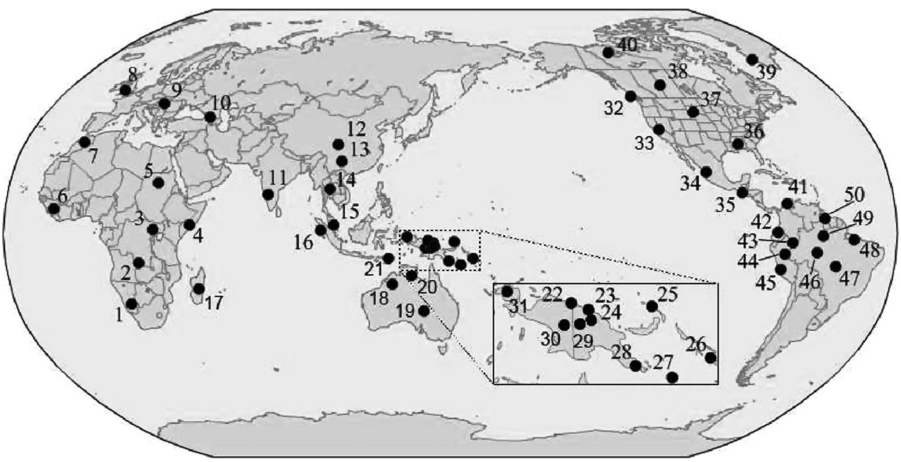


FIGURE 9.1 Sample languages on a world map (generated with a tool developed by Hans-Jörg Bibiko for *the WALS*). 1 Khoekhoe, 2 Lunda, 3 Ngiti, 4 Somali, 5 Dongolese (Nubia), 6 Kisi, 7 Middle Atlas Berber, 8 Welsh, 9 Hungarian, 10 Georgian, 11 Kannada, 12 Qiang, 13 Hmong Daw, 14 Thai, 15 Semelai, 16 Nias, 17 Malagasy, 18 Gooniyandi, 19 Diyari, 20 Alawa, 21 Klon, 22 Skou, 23 Arapesh, 24 Yimas, 25 Kuot, 26 Lavukaleve, 27 Yelî Dnye, 28 Daga, 29 Oksapmin, 30 Korowai, 31 Maybrat, 32 Nuuchahnulth, 33 Southern Sierra Miwok, 34 Cora, 35 Tzutujil, 36 Choctaw, 37 Lakhota, 38 Plains Cree, 39 West Greenlandic, 40 Slave, 41 Ika, 42 Imbabura Quechua, 43 Yagua, 44 Shipibo-Konibo, 45 Jaqaru, 46 Pirahã, 47 Trumai, 48 Urubú-Kaapor, 49 Hixkaryana, 50 Berbice Dutch Creole.

[4] Where feasible, languages were sampled randomly within genera, but in the case of isolates this was impossible.

9.3.3 *The experiments*

In the following subsections I present the results of the experiments, each dealing with a different characterization of complexity.

9.3.3.1 *Complexity as overt coding* Recall that my first hypothesis to be tested states that there is a complexity trade-off between the overt coding of one coding device and the absence of overt coding of another coding device in core argument marking. What this hypothesis predicts is that if a language has overt case marking, it will not have rigid word order and if a language has rigid order it will not have overt case marking.

In my sample, 3 languages use neither overt case marking nor rigid word order, 9 use case marking and rigid word order, 18 use only rigid word order, and 20 use only overt case marking. Figure 9.2 provides a summarizing heat map of the data.

I tested the hypothesis statistically by calculating Kendall's tau correlation coefficient. The reason for choosing a rank correlation test was that as a non-parametric test it makes few assumptions about the data and is thus well-suited for typological research, as argued by Janssen *et al.* (2006).[5] As suggested by the heat map in Figure 9.2, there are more languages that use only one coding device than those that use neither coding device or both of them. Such a distribution suggests that there is evidence for a trade-off. The correlation test confirms this observation: The variables correlated negatively and significantly with one another (Kendall's tau-a = −0.54;

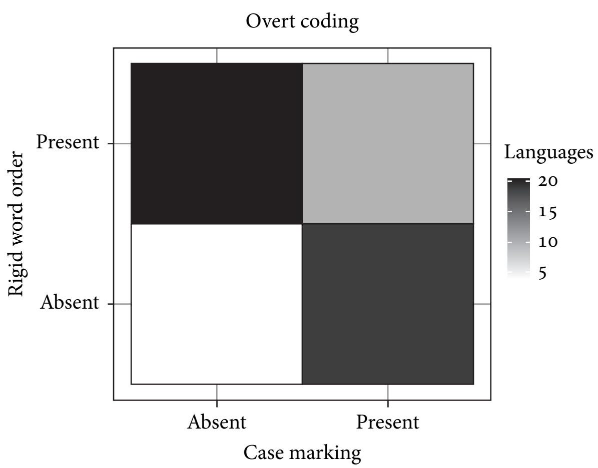


FIGURE 9.2 Heat map of complexity as overt coding. (The darkness of the shade reflects the number of languages in the particular cell.)

[5] All statistical tests and graphs were computed in the open-source statistical environment R (R Development Core Team 2012) and with additional packages *ggplot2* (Wickham 2009) and *rms* (Harrell 2012).

$p = 0.0002$).[6] Based on this result, there is empirical evidence for the trade-off hypothesis, namely, that overt case marking is balanced out by non-rigid order and lack of case marking is balanced out by rigid order. However, since 12 languages out of 50 violate the hypothesis (using both of the coding devices or neither of them), the complexity trade-off appears to be a statistical trend but not an absolute universal.

9.3.3.2 *Complexity as variety of marking patterns* My second hypothesis states that there is a complexity trade-off in core argument marking between a large variety of marking patterns in one coding device and a small variety of marking patterns in another coding device. What this hypothesis predicts is that if a language has a great variety of marking patterns (that is, many different forms) in case marking, it will have less variety of marking patterns (that is, fewer patterns) in rigid word order, and vice versa. Table 9.3 gives the language-specific data and Figure 9.3 a summarizing heat map.

A few words are needed to explain the diagnosis of case marking patterns in some languages. Alawa (Sharpe 1972) uses five patterns: singular A is marked with the ergative case, but optional gender marking has separate forms for the ergative (masc: *a-*; fem: *ar̄*) and absolutive arguments (masc: *na-*; fem: *an-*). Choctaw (Broadwell 2006) uses four patterns: both arguments are overtly case marked, but alternative forms are used for contrastive focus. In Diyari (Austin 1981) the A is marked with the ergative case and the P with the absolutive or the accusative case, as in Table 9.4. Since I exclude syncretism and zero forms, there are four different overt forms in the case marking of Diyari.

In Georgian both arguments are marked, but the particular case marker varies according to the tense of the verb, as in (7) (from Harris 1981: 1, glossing adapted from Palmer 1994: 80). Three different marking patterns are counted for Georgian, since syncretism is excluded.

(7) a. Glex-i tesavs simind-s.
peasant-nom sow.3sbj.3obj.prs corn-dat
'The peasant is sowing/sows the corn.'

b. Glex-ma datesa simind-i.
peasant-erg sow.3sbj.3obj.aor corn-nom
'The peasant sowed the corn.'

[6] I further double-checked whether the relationship between case marking and rigid order was independent of area by modelling the data with multiple logistic regression. Case marking and rigid order were modelled as binomial variables and area as a multinomial variable using the three-way areal breakdown into the Old World, the Pacific, and the New World (Nichols 1992). I chose rigid order as the dependent variable and case marking and area as the predictors. The effect of each variable was estimated using penalized maximum likelihood ratio to nested models and drawing the p-values by Monte Carlo randomized permutation testing (Baayen 2008: 205; Bickel 2008). A script developed by Bickel (2008) was used for the randomization. According to the results, the interaction term was non-significant ($LR = 0.1$; $df = 2$; $p = 0.95$) and so was the main effect of area ($LR = 3.3$; $df = 2$; $p = 0.18$). However, the main effect of case marking was very significant ($LR = 18.1$; $df = 1$; $p = 0.0001$). Because the interaction term was non-significant, the results suggest that the significant relationship between case marking and word order was independent of area.

Table 9.3. Variety of marking patterns. (The column 'Case marking' lists the case-marked argument and, when the number of forms cannot be inferred from this list, the number of different marking patterns used as well. The column 'Rigid order' lists the rigid order patterns used. The value is 'None' if the coding device is not used in the language.)

Language	Case marking	Rigid order
Alawa	A (five forms)	None
Arapesh	None	AVP, PAV
Berber (Middle Atlas)	A	VAP, AVP
Berbice Dutch Creole	None	AVP, PAV
Choctaw	A and P (four forms)	None
Cora	None	VAP
Cree (Plains)	None	None
Daga	None	None
Diyari	A and P (four forms)	None
Georgian	A and P (three forms)	None
Gooniyandi	A	None
Greenlandic (West)	A and P	APV
Hixkaryana	None	PVA, APV
Hmong Daw	None	AVP, PAV
Hungarian	P	None
Ika	A	None
Jaqaru	A and P (five forms)	None
Kannada	P	APV
Khoekhoe	P	None
Kisi	None	AVP, AauxPV, PAV
Klon	None	APV, AVP
Korowai	None	APV
Kuot	None	VAP
Lakhota	None	APV
Lavukaleve	None	APV
Lunda	None	AVP
Malagasy	P	VPA, VAP
Maybrat	None	AVP
Miwok (Southern Sierra)	A and P	None
Ngiti	None	AVP, AauxPV
Nias	P	VPA
Nuuchahnulth	None	VAP
Nubian (Dongolese)	P	None
Oksapmin	P (two forms)	None
Pirahã	None	APV
Qiang	A and P	None
Quechua (Imbabura)	P	None
Semelai	A and P	AVP
Shipibo-Konibo	A	None
Skou	A	APV

Slave	None	APV
Somali	A (four forms)	None
Thai	None	AVP
Trumai	A	APV, PAV
Tzutujil	None	VPA
Urubú-Kaapor	P	None
Welsh	A	VAP, AVP
Yagua	None	VAP, AVP, PVA
Yelî Dnye	A (two forms)	None
Yimas	None	None

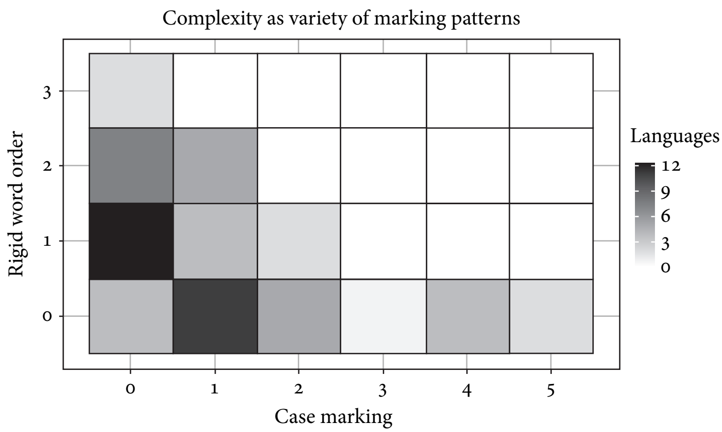


FIGURE 9.3 Heat map of complexity as variety of marking patterns.

c. Glex-s dautesavs simind-i.
peasant-dat sowed.3sbj.3iobj.pfv corn-nom
'The peasant has sown the corn.'

Jaqaru (Hardman 1966, 2000) uses five case marking patterns, as explained in §9.2.2.2. Oksapmin (Loughnane 2009) uses two patterns: The P can be case marked for higher animates, using *=nuŋ* with a pronominal article and *=ja* when there is no pronominal article (the articles indicate specificity). Somali (Saeed 1999) uses four patterns. The accusative is in the citation form, while the nominative is marked with low tones on all moras, except that non-plural feminine nouns ending in *-o* and masculine nouns ending in *-i* use a high tone on the penultimate mora. Furthermore, feminine nouns ending in a consonant are marked with *-i* in the nominative and definite suffixes have a separate form for the nominative. Yelî Dnye (Henderson 1995) has two patterns: The A is case marked with *ngê* for singular and *yyoo* for plural arguments.

TABLE 9.4. **Case marking forms in Diyari (Austin 1981: 47–49)**

	Ergative	Absolutive	Accusative
Proper nouns			
Male personal names	-li	-n̪a	
Female personal names	-ndu		-n̪a
Common nouns			
Singular	-yali	zero	
Non-singular	-li		-n̪a

The heat map in Figure 9.3 suggests that languages with a lot of variety in rigid order tend to have less variety in case marking and vice versa. This distribution indicates that there is evidence for a trade-off, which is confirmed by the correlation test: The variables correlated negatively and significantly with one another (Kendall's tau-a= -0.48; $p < 0.0001$), although in a slightly weaker way than when examining complexity as overt coding (tau-a = -0.54). The data suggest, therefore, that there is evidence for the trade-off hypothesis, namely, that languages with a great variety of marking patterns in case marking tend to have less variety of marking patterns in rigid order (and vice versa).

9.3.3.3 *Complexity as system dependencies* My third hypothesis states that there is a complexity trade-off between a large number of system dependencies in one coding device and a small number of system dependencies in the other coding device. What this hypothesis predicts is that if case marking depends on a large number of grammatical systems in a language, rigid order depends on fewer grammatical systems and vice versa. Table 9.5 gives the language-specific data and Figure 9.4 a summarizing heat map of the data.

The heat map in Figure 9.4 suggests that there is some tendency for languages in which case marking depends on many grammatical systems to have fewer dependencies in rigid order and vice versa. However, there is also a large group of languages in which neither coding device depends on other grammatical systems (the darkest cell in the figure). Such a distribution suggests that there is a negative correlation between the variables, but its strength is questionable. The correlation test in fact confirms that the variables correlated negatively but not significantly with one another (Kendall's tau-a = -0.18; $p = 0.18$). This outcome may be explained by the fact that it requires the same description length (that is, zero) to account for the system dependencies in a language in which overt case marking does not depend on other grammatical systems as in a language in which there is no overt case marking at all, since in such languages case marking does not depend on other systems (and similarly for rigid order). Based on these results, the data provide evidence for a

TABLE 9.5. **System dependencies of case marking and rigid order. (The columns 'Case marking' and 'Rigid order' list the grammatical systems that the particular coding device depends on. The value is 'None' if there are no dependencies and 'Unknown' (counted as one) if there is variation in the marking patterns, but the dependency could not be identified from the sources. 'Noun class' refers to the distinction proper/common nouns.)**

Language	Case marking	Rigid order
Alawa	Gender, number	None
Arapesh	None	Information structure (topicalization)
Berber (Middle Atlas)	Noun class, information structure (topicalization)	Information structure (topicalization)
Berbice Dutch Creole	None	Information structure (focusing)
Choctaw	Information structure (focusing)	None
Cora	None	None
Cree (Plains)	None	None
Daga	None	None
Diyari	Gender, number, noun class	None
Georgian	Tense	None
Gooniyandi	Animacy	None
Greenlandic (West)	Number	Number
Hixkaryana	None	Information structure (emphasis)
Hmong Daw	None	Information structure (topicalization)
Hungarian	None	None
Ika	(Unknown)	None
Jaqaru	Animacy	None
Kannada	Animacy, number, information structure (focusing)	Animacy
Khoekhoe	Information structure (focusing)	None
Kisi	None	Information structure (focusing), Tense
Klon	None	Animacy, information structure
Korowai	None	None
Kuot	None	Information structure (topicalization)
Lakhota	None	Animacy
Lavukaleve	None	None
Lunda	None	None
Malagasy	Noun class	Information structure (topicality)
Maybrat	None	None
Miwok (Southern Sierra)	None	None
Ngiti	None	(Unknown)

(*continued*)

Table 9.5. Continued

Nias	Aspect	Aspect
Nuuchahnulth	Animacy	Animacy
Nubian (Dongolese)	Animacy, definiteness	None
Oksapmin	Animacy, specificity	None
Pirahã	None	None
Qiang	Animacy, information structure (topicalization)	None
Quechua (Imbabura)	None	None
Semelai	Animacy, (unknown)	Non-individuated event
Shipibo-Konibo	None	None
Skou	(Unknown)	None
Slave	None	None
Somali	Gender, number, definiteness	None
Thai	None	Information structure (topicalization)
Trumai	None	Information structure (focusing)
Tzutujil	None	Definiteness
Urubú-Kaapor	Animacy	None
Welsh	None	Aspect
Yagua	None	Information structure (focusing)
Yelî Dnye	Number	None
Yimas	None	None

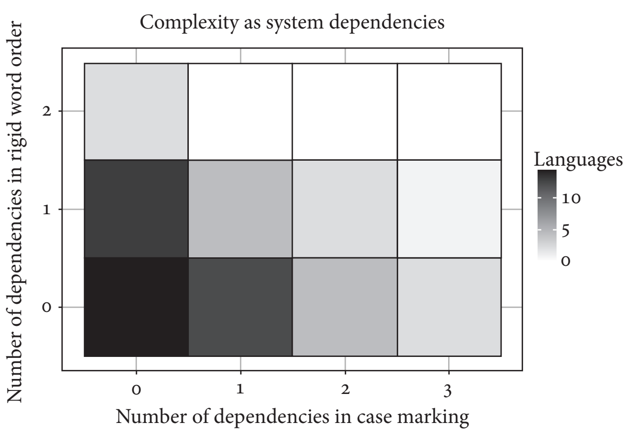


Figure 9.4 Heat map of complexity as system dependencies.

complexity trade-off between case marking and rigid order in terms of resources, but not regulations.

9.4 Discussion

We have seen that there is a significant negative correlation between case marking and rigid word order with respect to overt coding and with respect to marking patterns, but not with respect to system dependencies. The data thus suggest that a complexity trade-off occurs between case marking and rigid order, but that trade-off is neither absolute nor all-encompassing: Instead, the trade-off is statistical and depends on the particular grammatical characteristic focused on, namely, describing case marking and rigid order as 'resources' (building blocks), not as 'regulations' (roughly 'rules of grammar'). These results raise at least the following questions, which I address briefly in what follows: 1) Why is there a complexity trade-off at all?; 2) Why is the trade-off not absolute?; 3) What possible processing mechanisms could cause the trade-off?; and 4) Why are resources but not regulations of the system important to the trade-off?

I treat the first three questions together, because they are interrelated. Languages which adhere to the trade-off hypothesis use only one coding device (either case marking or rigid word order), or they have less variety in the marking patterns of one coding device compared to that of the other coding device. Using a single coding device instead of two, or mutually limiting the variety of marking patterns of the coding devices, is a manifestation of the Zipfian principle of least effort or economy (e.g., Hawkins 2004; Hornstein *et al.* 2005; Haspelmath 2008). As a processing principle, however, economy is not merely a matter of 'minimize all' but rather 'minimize all you can', that is, dispense with all unnecessary distinctions (Bornkessel-Schlesewsky and Schlesewsky 2009). Essentially, a system which arrives at distinctness in an economic way is likely to be easier in language acquisition and language use compared to a system which violates economy (see especially Hawkins 2004). In that sense, formal complexity may be affected by processing principles—and thus usage-based complexity.

But why do some sample languages violate the trade-off hypothesis? In the case of those languages which use no case marking or rigid order (Daga, Plains Cree, and Yimas), the answer is relatively simple: They use agreement instead, as in Daga (8) (Murane 1974: 136). Using agreement for argument discrimination in these languages violates neither distinctness nor economy in the argument marking system as a whole.

(8)	Apan	yampo	orup oaen	den	mokare	yaw-an-e.	(Daga)
	man	three	girl	with	snake	see-3pl.sbj.pst-med	

'The three men and the girl saw a snake . . . '

A departure from economy occurs in the sample languages which use both case marking and rigid order, since they arrive at distinctness in uneconomic ways. In five of these languages (Malagasy, Middle Atlas Berber, Skou, Trumai, and Welsh), the two coding devices are used in the same constructions, which clearly violates economy. The reasons for these violations may have to do with the diachrony of each language. However, in the four other languages (Kannada, Nias, Semelai, and West Greenlandic), the coding devices are used in roughly complementary ways. For instance, in Semelai the arguments are discriminated by case marking (but not rigid order) in the perfective clauses, but by rigid order in clauses which express generic meaning (Kruspe 1999). While the overall system of core argument marking is perhaps unnecessarily complex in these languages, the complementarity of case marking and rigid order demonstrates that in the particular constructions the languages arrive at distinctness via economic means.

As for the fourth question, the results suggest that there is an interaction between case marking and rigid order with respect to resources, rather than regulations, in the system. It seems that complexity trade-offs in other grammatical areas as well might be possible only in terms of resources but not regulations. In comparing the complexity of rules in phonology, Parker (2009) found no interesting trade-offs at all. Rather, there seemed to be just random variation. Fenk-Oczlon and Fenk (1999), on the other hand, found complexity trade-offs between different phonological variables when measuring, for instance, word length and sentence length in terms of the number of phonemes and syllables. Thus, while the focus on building blocks (resources) as central to complexity measurement has been severely criticized (e.g. Arends 2001; DeGraff 2001), there is some typological evidence that complexity trade-offs are more relevant to resources than to regulations in the system. But why should this be so? One reason may be that there is a more general tendency for interactions in grammar to occur between the parts of the system, rather than between interactions in the system. However, we must leave this question as a matter for further research.

9.5 Conclusion

I have argued here that the general notion of complexity can be meaningfully approached in cross-linguistic research. This approach focuses on local complexity, separates complexity from language use, and measures complexity as description length. Given such a focus, I was able to demonstrate in a case study that a strong complexity trade-off occurs between case marking and rigid order in core argument marking. Interestingly, the trade-off depended on the way grammar was described: It only occurred when focusing on resources, or building blocks, in the system, but not when focusing on regulations, or 'rules of grammar', in the system. Moreover, the mere presence or absence of the coding devices was relevant to the trade-off. These results suggest that the question of complexity trade-offs is complex in itself and that even rather crude analytic tools may provide feasible starting points for measuring grammatical complexity.

10

The importance of exhaustive description in measuring linguistic complexity: the case of English *try and* pseudocoordination

DANIEL ROSS

10.1 Introduction

For practical considerations such as time and available data, studies attempting to measure linguistic complexity typically rely on subsets of the properties of a language. However, it is the argument of the current chapter that such measurements are likely to be inaccurate. We will see that properties that appear to be exceptional, infrequent, obscure or marginal, are not well documented, or are otherwise peripheral in a grammatical system can and do increase complexity. Section 10.2 discusses linguistic complexity at a general level, while section 10.3 considers specific examples relevant to complexity measurement, with a particular focus on English *try and* pseudocoordination. Section 10.4 is a brief summary and discussion.[1]

10.2 Grammars and complexity

I adopt here an essentially length-based definition of complexity. That is, if grammar A and grammar B are otherwise identical, yet grammar B has a rule, construction, principle, or whatever that is missing in grammar A, then B is ipso facto more complex than A. I begin below by discussing how a grammar can be characterized for

[1] I wish to thank Abbas Benmamoun, Frederick Newmeyer and Laurel Preston for their detailed comments on earlier drafts of this chapter, as well as the participants of the *Formal Linguistics and the Measurement of Grammatical Complexity Workshop*, March 2012, at the University of Washington for their feedback.

the purpose of measuring complexity, and then return to this length-based definition of complexity in section 10.2.2.

10.2.1. *Characterizing grammatical systems for complexity measurement*

The crucial question is precisely what we mean when we claim that a particular language has such-and-such a degree of complexity or that one grammatical system is more complex than another. Before we can begin to answer such questions, we need to decide precisely what we are going to measure and what grammatical theory we are going to adopt on which to base the measure. As far as the first point is concerned, it needs to be stressed that the very notion of 'language' is a socio-political abstraction: grammars rest in the minds of individuals and not all speakers of 'English' have identical grammars. In principle, then, I conceive of a complexity measure as pertaining to individuals, not groups. As is commonly the case in grammatical research, however, I abstract away from the individual–group distinction, hoping that the consequences of doing so will not be negative. More importantly and more controversially, we need to know what to measure when we are measuring complexity. For example, we could base our measure on the relative ease with which speakers acquire grammatical knowledge; we could base it on the relative ease with which speakers use this knowledge; or we could base our measure on the complexity of grammars taken as 'objects', abstracting away from acquisition and use. This chapter adopts the latter option, namely a 'grammar-based' as opposed to a 'user-based' definition of complexity.

But there are many different theories of grammar and there is no reason to think that they would all lead to the same complexity measurement. Is a grammar constructed with both phrase-building rules and displacement rules more complex or less complex than one with phrase-building rules alone? Needless to say, when comparing grammars of two different languages in terms of their relative complexity, it is important to ensure that they have been characterized using the same theoretical model. This chapter remains agnostic on which model of grammatical theory is the 'best'. I believe that the phenomena that I discuss below, in particular *try and* pseudocoordination, are likely candidates for increasing complexity regardless of the particular framework used to describe them. Finally, I need to stress that the data considered here are almost entirely morphosyntactic in nature. I have every reason to believe, however, that my general approach would extend to other grammatical components.

10.2.2. *Defining complexity*

The length-based approach to complexity mentioned above has been applied by others in measuring grammatical complexity:

> One may think of grammatical complexity in terms of a textbook of grammar. Roughly speaking, the longer the textbook, the more complex the grammar. This agrees very well with the notion of complexity as the length of a schema. Every nasty little exception adds to the length of the book and the grammatical complexity of the language. (Gell-Mann 1994: 55)

Known as 'Kolmogorov complexity' or 'Algorithmic Information Content', this approach crudely predicts that the most complex objects are those with completely *stochastic* (in linguistic terms, *arbitrary*) content, without any internal patterns. In other words, the most complex object would be one in which the shortest possible description of an object would be the object itself (cf. Gell-Mann 1994; Dahl 2009: 50–51).

In language, much arbitrary content is stored in the lexicon. I choose here to ignore properties of the lexicon in measuring complexity. I agree completely with Östen Dahl when he writes:

> One often-heard reason for skepticism about using complexity measures that build on length of descriptions is that they seemingly make a language with a large vocabulary more complex than one with a small vocabulary. (Dahl 2009: 52)

If we were to include the lexicon in our complexity measurement, then, based on vocabulary size, English would be the most complex language in the world (cf. McCrum, Cran, and MacNeil 1992: 1), although that is of little theoretical interest. Furthermore, counting the number of words would represent the *theoretical* lexicon of English, rather than the knowledge of the typical native speaker. I therefore adopt what Gell-Mann calls 'effective complexity' (cf. Gell-Mann 1994: 58), which would allow us to ignore information in the lexicon and consider only patterns in the grammar.

There remains the question of whether lexical instantiations of grammatical patterns contribute to complexity. Consider the present-tense verbal paradigms of English and German in Figure 10.1.

English present tense		
	Singular	Plural
1st person	*play*	*play*
2nd person	*play*	*play*
3rd person	*plays*	*play*

German present tense		
	Singular	Plural
1st person	*spiele*	*spielen*
2nd person	*spielst*	*spielt*
3rd person	*spielt*	*spielen*

FIGURE 10.1 English and German present-tense verbal morphological paradigms.

The German paradigm appears more complex than that of English because there are more morphological forms and therefore, apparently, more properties. Indeed, the number of morphological forms has been considered relevant in previous research measuring morphological complexity (e.g. Dahl 2009; McWhorter 2001, *inter alia*). But my position is that this difference between English and German resides in the lexicon.[2] According to standard grammatical analyses, verbal agreement in both languages is based on two underlying features, *person* and *number*. In this way, English and German contrast with other languages with fewer distinctions (such as Swedish with no subject agreement) or additional contrasts (such as *verbal gender* in Arabic). The differences between English and German are therefore lexical in nature. In German, there are additional suffixes realizing certain person/number combinations, while in English only third-person singular has a unique morpheme in the lexicon. For grammatical complexity, there is no difference between the systems, since both involve subject-verb agreement and in both the features of person and number are contrastive.[3] To measure grammatical complexity, we focus on whichever linguistic properties are active in the grammar, not how these properties are realized lexically.

In summary, the length of a grammatical description (excluding the lexicon) is one reasonable definition of grammatical complexity, and the one that will be used here. Crucially, all measurable properties contribute to overall complexity, and hence none can be ignored. The remainder of this chapter is devoted primarily to examining what one might consider to be ‘peripheral’ properties of a grammar, showing how they increase grammatical complexity.

10.2.3 *The importance of exhaustive description*

Given that our goal is to measure the knowledge of a speaker, all aspects of the grammar of a language must enter into an adequate complexity measure, including those that might for whatever reason be considered ‘marginal’ or ‘peripheral’ (but see Sinnemäki, this volume, chapter 9 for an opposing viewpoint). Unfortunately, many studies make claims about relative complexity that seem to focus on readily apparent and easily documented properties of a language. One example is McWhorter (2001), where it is argued that Saramaccan, a creole, is simpler than older languages like Tsez,

[2] Paradigm size can affect processing difficulty, but here the focus is on grammatical complexity, rather than processing. No claim is made that complexity and processing difficulty are necessarily correlated; in fact, it would not be surprising if two grammatical properties of roughly equal grammatical complexity turned out to be significantly different in processing difficulty.

[3] We could bow to the popular claim that paradigm complexity is relevant to complexity, which, given standard analyses, would make German more complex than English in this respect. Alternatively we could entertain a different analysis available for English, say, employing the binary feature [±3SG], rather than the two features of *person* and *number*. I will leave this as an open question, but proceed with the analysis presented above.

Lahu, and Maori. Claiming in the title that "the world's simplest grammars are creole grammars," McWhorter generalizes beyond what his data support. His analysis is based on a small set of hand-picked features, which he takes to be representative. Such work is problematic because he gives us no reason to believe that other undiscussed aspects of Saramaccan are necessarily simpler than other undiscussed aspects of Tsez, Lahu, and Maori.

There are several reasons why a property might be neglected in a measurement of complexity. One possibility is that a particular phenomenon is not understood well enough to be included due to, say, sketchy data gathered from an understudied or moribund language in fieldwork. Or the phenomenon might be omitted due to difficulty integrating it into a grammatical theory. Another reason that a property might be overlooked could be due to its appearing only in low frequency constructions. We know from Zipf (1935) that the distribution of lexical items is not linear but logarithmic. For example, about 300 words forms make up 50% of the tokens in the 450 million word Corpus of Contemporary American English (COCA; Davies 2008–; data from http://www.wordfrequency.info). The same applies to the distribution of syntagms in English, as reported by Köhler (2007: 195). Fewer than 500 constituent types account for 90% of the data in a corpus with a total of about 4500 constituent types (Beebe 1973 presents similar results). It is therefore likely that many infrequent grammatical properties have been overlooked simply on the basis of their rarity. However, as Dahl observes, the frequency of individual forms "would not influence the complexity of the grammar" (2009: 59), because regardless of frequency in performance, the knowledge required to produce these forms is part of the speaker's competence in the language.

10.3 Peripheral phenomena and complexity measurement

This section provides an overview of the type of linguistic phenomena that should not be neglected in measurements of complexity. First, I discuss examples from several languages with respect to their contribution to complexity. Second, I present a case study of *try and* pseudocoordination, and demonstrate that it adds to the grammatical complexity of English. The section concludes with a brief discussion of why pseudocoordination in other languages, in contrast to English, does not necessarily increase complexity.

10.3.1 *Searching for complexity crosslinguistically*

Let us examine some crosslinguistic examples of how overall complexity is increased by phenomena that one might characterize as 'peripheral', 'marginal', or 'less salient'. To begin, consider Spanish, which is commonly thought of as having two genders: masculine and feminine. That would make Spanish more complex than English (with no grammatical gender), but less complex than German (with three grammatical genders). But in fact, in one analysis, Spanish equals German in its number of genders and hence (in this respect) in its degree of complexity. Spanish has traces of the neuter gender from Latin, as shown in a handful of function words: the demonstratives *eso*, *esto* and *aquello*, the article *lo* used in the expression *lo que* ('that which') or with adjectives (*lo importante*, 'the important [thing]'), and the archaic pronoun *ello* ('it'). The Spanish grammatical system must have some means by which to distinguish these forms from other masculine and feminine forms, if not by a third gender, then at least by some other complexity-increasing device.

In Swahili, a Bantu language, the approximately 15 noun classes usually pattern regularly for agreement, but there are some unexpected exceptions (cf. Mohammed 2001: 47–51). So in example (1), *rafiki*, which appears to belong to Class 5 based on the form of the plural *marafiki* and Class 5 possessive agreement in *yangu*, requires Class 1 subject–verb agreement, which is appropriate for animate subjects:

(1) Rafiki yangu anapenda kucheza futboli.
friend my.CLASS5 likes.CLASS1 to.play football
'My friend likes to play soccer.'

Since there is no one-to-one correspondence between lexical noun class and agreement or plural formation, the combinatorial possibilities and therefore the complexity of Swahili are increased beyond what we would expect from a simple, regular agreement system. Examples like these are common crosslinguistically.

At the same time, certain properties that superficially appear to add complexity to a language may be explainable by existing, more general, properties and therefore do not result in an increase in complexity. One example is the definite marker in Amharic (see Kramer 2010). On the surface, the distribution of the definite marker appears to be quite complex. As shown in (2–3), it always appears at the end of the first constituent within the noun phrase:

(2) k'ondʒo-**w** tɨllɨk' k'äyy kwas
beautiful-DEF big red ball
'the beautiful big red ball' (Kramer 2010: 200)

(3) ɨdʒdʒɨg bät'am tɨllɨk'-**u** bet
really very big-DEF house
'really very big house' (Kramer 2010: 198)

Despite the apparently complex distribution of the definite marker, Kramer explains its properties with reference to general mechanisms within the framework of Minimalism and Distributive Morphology. One of the lexical properties of the definite marker *-u* is that it is an enclitic and therefore must attach to the end of a constituent. This fact can be explained by Local Dislocation, which allows the linear inversion of two phonological forms if one must attach to a certain edge of the other. Hence we can analyze the definite marker as the first node in the noun phrase, which is then displaced to the end of the first constituent. The definite marker attaches to the end of the first full constituent, not the first word, because of Cyclic Spell-out and Phase Impenetrability, which states, essentially, that part of the derivation that is already completed (spelled-out) is no longer accessible to the rest of the derivation except as a whole. Therefore, the enclitic *-u* attaches to the closest available location within the noun phrase, which happens to be the end of the first constituent. In this way, what appears to be a complex aspect of Amharic turns out to be no more complex than definiteness marking in a language like English, because the distribution is readily explained by more general properties of the language. In other words, the complexity of any individual component of a language must be measured in the context of the language as a whole.

10.3.2 *Why* try and *pseudocoordination makes English more complex*

Verbal pseudocoordination, as found in the phrases *go and get* or *try and do* in English, is relatively understudied outside of the Germanic languages of Scandinavia (for which, cf. Wiklund 2007). This section focuses on the *try and* type of pseudocoordination in English, and argues that its analysis requires appeal to a 'bare form condition' (Carden and Pesetsky 1977). Since this condition cannot be derived from independently motivated principles, *try and* pseudocoordination therefore adds to the complexity of English.

Try and pseudocoordination has caught the attention of a number of researchers, some of whom are listed below, but to date no adequate account has been published of this construction. Previous accounts have primarily noted the construction as prescriptively incorrect or illogical (Waddy 1889: 110–111; Poutsma 1917), or as an unexplained exception (Gleitman 1965: 293; Ross 1967). Others have discussed the descriptive generalizations in detail but do not attempt a theory-based explanation (Carden and Pesetsky, 1977; Pullum 1990), while corpus research has established usage patterns (Lind 1983; Hommerberg and Tottie 2007). The most comprehensive published accounts are presented by Hargreaves (2005) and de Vos (2005: 202–207), based on Hargreaves (2004), but even these do not address certain relevant properties discussed below.[4]

To begin the analysis (which is presented in more detail in Ross 2013a), we must distinguish pseudocoordination from standard coordination. First note that the

[4] I cannot provide a full description of these accounts here, but for several reasons they do not contradict the claims made below. In the first place, these accounts are incomplete; de Vos (2005: 203) explicitly presents the argument as a suggestion for further research, while Hargreaves (2005: 32) does not, for example, provide a theoretical explanation for the ungrammaticality of (31) below. Further, it appears that for these arguments to be successful they would still require some explicit reference to inflection (de Vos 2005: 203), as in the Inflectional Parallelism Requirement to be presented here.

former is different from the latter in often having a purposive meaning (much like *to*-infinitives). So one possibility to be considered is that pseudocoordination is at root standard coordination, but with pragmatic implicatures that extend the meaning of coordination to purpose. Such is very likely the origin of such constructions, but pragmatics alone cannot account for the unique grammatical properties outlined below. As shown in (4–5), the word order of the conjuncts is fixed in pseudocoordination, and as shown in (6–7), the pseudocoordination reading is restricted to verbal conjuncts without intervening subjects. Also unlike standard coordination, pseudocoordination is strictly binary, as shown in (8–9).

(4) I try and do it. ≠ I do and try it.
(5) I go and get it. ≠ I get it and go.
(6) I try and do it. ≠ I try and I do it.
(7) I go and get it. ≠ I go and I get it.
(8) *I try, go and get it.
(9) *John took the book, up and stole it.

Also note that pseudocoordination allows violations of the Coordinate Structure Constraint (CSC; Ross 1967: 167–170): Unbalanced extraction out of only one conjunct is possible (10–11), unlike in standard coordination (12):[5]

(10) What did you try and do?
(11) What did you go and get?
(12) a. *What did you {rest / read a book} and want?
b. *What kind of food did you drink (a soda) and eat?

The most distinctive property of pseudocoordination may be that it has a strict requirement of parallel morphology on both verbs (see Wiklund 2007 for extensive discussion of this phenomenon), as shown in (13–14):

(13) *I tried and do it.
(14) *I went and have a book.

Having established that pseudocoordination is distinct from standard coordination, let us now distinguish two subtypes of pseudocoordination in English. The first, the *go*-type (e.g., *go and V*), is found most often with motion verbs, and it is common crosslinguistically, as discussed in section 10.3.3 below. This type has existed since at least the Old English period, when it was often used to translate Latin purposive infinitives (Shearin 1903: 12–13). The second, the *try*-type (e.g. *try and V*), appears to be unique to English (but see the discussion of Faroese below). After the

[5] However, Lakoff (1986) argues that in general the CSC is only a semantic constraint and that even long logical sequences of coordinated actions allow unbalanced extraction, as in (i):

i) Sam is not the sort of guy you can just sit there, listen to, and stay calm. (Lakoff 1986: 153) lexical item *trie* was borrowed from French in the twelfth century, with the original meaning of 'test' or 'examine' (OED: *try*, v.), both *try and* and *try to* developed about the same time with the meaning of 'attempt' in the late 1500s (Hommerberg and Tottie 2007: 60; Ross 2013b: 115).

Although similar, the two types of pseudocoordination differ semantically. For example, the *try*-type does not entail the completion of the action described by the second verb (15), while the *go*-type does (16), and the most appropriate paraphrases suggest opposite directions of semantic subordination (17–18), with the second verb acting as an argument of the first for the *try*-type in (17) and the first acting as a modifier of the second for the *go*-type in (18):

(15) I will try and finish the report on time, but I might not succeed.
(16) I will go and get the book (#even if it is sold out).
(17) I try <u>and get</u> good grades. ≈ I try <u>to get</u> good grades.
(18) I <u>go and</u> get the book. ≈ I get the book <u>by going</u>.

Grammatically, the *try*-type is subject to the *bare form condition* (Carden and Pesetsky 1977), which restricts inflection (19–20). This condition does not apply to the *go*-type (21–22).

(19) *He tries and do(es) it.
(20) *We tried and do/did it.
(21) He goes and gets it.
(22) He went and got it.

To show that this condition is sensitive to inflection rather than, for example, semantics, we can use *be sure and*. The uninflected future form, for example, is grammatical (23); in the present tense, however, *be* is not licensed by pseudocoordination nor its syntactic context (24), and inflected *am* violates the bare form condition (25).

(23) I will be sure and take out the trash.
(24) *I be sure and take out the trash.
(25) *I am sure and take out the trash.

Likewise, the bare form condition also holds for the second verb (26). However, notice that the *try*-type licenses *be* as the second verb in the present tense (27), despite the syntactic context.

(26) *I try and am happy.
(27) I try and be happy.

Although most common with *try* and *be sure*, *try*-type pseudocoordination is also found marginally (infrequently, and to varying degrees of acceptability for different speakers) with other first verbs including *pretend*, *prefer*, *remember*, *promise*, *manage*, and to some extent even infrequent verbs like *yearn* (see Ross 2013a). Examples can be found in corpora (28), and such usage is accepted by some speakers in acceptability judgment tasks (29):

(28) Remember and wash your hair. (British National Corpus: KE4 636)
(29) Sam likes to pretend and do his homework. (Ross 2013a: 19)

The class of verbs that can participate in the *try*-type appears to be limited to control verbs, specifically only those that have obligatory subject control. This excludes object-control verbs (**ask and*...), ECM verbs (**want and*...), and raising verbs (**seem and*...). Hence *try*-type pseudocoordination is a grammatical phenomenon, not a lexical idiom found only with *try*.[6]

In what follows I will show that the bare form condition is a consequence of two separate grammatical principles, namely agreement and parallelism. First, let us consider the nature of both verbs in the construction. In the present tense *be* is licensed as the second verb as in (27) above, but not the first as in (24). The second verb is not required to agree with the subject, and thus must be a bare infinitive, licensed by *and* (in a manner at least superficially similar to complementation with *to*-infinitives). Given this fact, the ungrammaticality of (30) follows naturally. In this manner, the half of the bare form condition that applies to the second verb has been derived.

(30) *He tries and sleeps.

Conversely, although the first verb looks like a bare infinitive, it must agree with the subject and otherwise be licensed in its syntactic position. That accounts for why (24) above and (31) below are ungrammatical:

(31) *He try and sleep.

At this point, with only subject-verb agreement for the first verb and a bare infinitive required for the second verb, the grammar would predict (32) to be grammatical, yet it is not, in contrast to the bare first-person present-tense (33):

(32) *He tries and sleep.
(33) I try and sleep.

Therefore, the second grammatical property of *try*-type pseudocoordination is that there is a parallelism requirement on the forms of the two verbs. I now argue that this property adds to the complexity of English, because it does not follow from anything more general in English grammar.

[6] However, for convenience and due to the relatively high frequency of *try* as the first verb in this construction, I will continue to use the label '*try*-type,' though the grammatical properties hold for the other verbs as well (but not those of the *go*-type).

As a starting point, it is widely known that coordination involves, in some sense, parallelism (cf. Munn 1993; te Velde 2005; *inter alia*). For example, Wasow's Generalization (Sag, Gazdar, Wasow, and Weisler 1985; Pullum and Zwicky 1986: 752 –753), explains the grammaticality of (34), because despite not being of the same type, both coordinated elements have appropriate syntactic features such that they would be independently licensed to appear in place of the full coordinate structure:

(34) He's a Republican and proud of it. (Pullum and Zwicky 1986: 752)

In the context of pseudocoordination, a requirement of parallel morphology has been observed crosslinguistically (see Wiklund 2007, *inter alia*). In terms of concrete analyses, this parallelism has been explained for the Scandinavian Germanic languages as *feature copying* (Lødrup 2002) or *tense agreement* (Wiklund 2007), which offer equivalent predictions of morphological parallelism at the level of syntactic features. Although—in line with these analyses—the *go*-type in English can be explained by means of coordinator-imposed featural agreement for the two verbs, the *try*-type is more difficult to explain. The morphological parallelism for English *try*-type pseudocoordination does not operate at the level of syntactic features: the combination of present-tense *try* and infinitive *be* is possible, as is illustrated in (27), repeated here as (35), in contrast to the *go*-type (36):

(35) I try and be happy.
(36) *I go and be happy.

One might argue that morphological parallelism for the *try*-type is a purely surface phenomenon. However, such cannot be the case, given the ungrammaticality of (37):

(37) *He tried and put the book on the table.

Therefore, the parallelism requirement must hold at the level of morphological inflections:

> Inflectional Parallelism Requirement: Every verbal complement in *try*-type pseudocoordination must manifest the inflection of the matrix verb.

This requirement replaces any need for a literal bare form condition: the second verb is necessarily bare because it is a bare infinitive, while the parallelism requirement makes realization of any overt inflection on the first verb impossible as well. It is the Inflectional Parallelism Requirement, with its explicit reference to morphological inflection, that adds to the complexity of English grammar.

One might speculate that parallelism of inflection could be reduced to the Resolution Principle (RP) of Pullum and Zwicky (1986), which was introduced to account for the contrast between examples like (38) and (39):

(38) *Either you or I am/are here.
(39) Either you or they are here.

The RP states that in the case of conflicting syntactic features on coordinated terms, as in (38) and (39) above, the conflict can be resolved if "a phonological form is available, which is, at the relevant [lexical?] stratum of representation, ambiguous between these values" (Pullum and Zwicky 1986: 766). In other words, (39) is grammatical because *are* is the same form for second-person and third-person plural. We might wonder, then, whether the bare present-tense first verb and bare infinitive second verb forms in (35) are analogously ambiguous at a certain level of representation. Let us assume such an analysis is possible; does this eliminate the need for additional complexity? Pullum and Zwicky present the RP as a universal (1986: 767). Perhaps we might speculate that it is even extra-grammatical, as a natural consequence of processing. While it would be appealing to derive *try*-type pseudocoordination from a universal principle and thereby avoid postulating reference to inflections in this syntactic construction, the RP is not universal: although present-day English offers no overt evidence to that effect, Faroese does. This language has a similar construction to the English *try*-type pseudocoordination, which is limited in distribution to non-finite contexts (Heycock and Petersen 2012: 274):

(40) Tú mást royna og lesa bókina! (Heycock and Petersen 2012: 275)
you must try.INF and read.INF book.the
'You must try and read the book.'

(41) *Eg royni og lesi bókina.
I try.1S and read.1S book.the
('I try and read the book.')

However, because the infinitive and plural present-tense forms are homophonous in Faroese, some speakers today colloquially accept plural present-tense usage of the *try*-type (Heycock and Petersen 2012: 274). Like English, Faroese only allows *try*-type pseudocoordination with non-finite forms, or other forms that are inflected similarly:

(42) Tey royna og lesa bókina.
They try.3PL and read.3PL/INF book.the
'They try and read the book.'

Because structures of this type are possible for some speakers, Faroese presents evidence against the universality of the RP, given that this principle would be violated. In fact, earlier English appears to have been similar to standard Faroese: while *try and* pseudocoordination is attested since the late 1500s (43), finite present-tense usage is not found until the 1800s (44):[7]

[7] See Ross (2013b: 119–120). Further support for this change in the 1800s is provided by reference guides from the time, such as Waddy (1889: 147 –148), who restricts the pseudocoordination reading to '*root-forms*' (infinitives, and presumably imperatives), while present-tense forms only have the standard coordination reading ('try, and [then] do').

(43) You maie (saide I) trie and bring him in, and shewe him to her. (Year: 1569; EEBO TCP Corpus)

(44) Do sit down by the fire, whilst I try and get you some breakfast. (Year: 1841; Google Ngrams: Michel *et al.* 2011)

The universality of the RP could be maintained if there were some other explanation for inflectional parallelism. Such is a logical possibility, but any explanation would still need to make reference to inflection-sensitive parallelism of the two verbs (or a literal bare form condition, stipulating that surface-level inflection is not possible), independently of syntactic features. For example, Construction Grammar might postulate a *try and* construction with an inflectional restriction, or Minimalism some version of *inflection-checking*. Regardless, any theory would have to make explicit reference to inflection, in order to differentiate the grammars for present-day English from earlier English, and to characterize the colloquial Faroese discussed above.

In summary, the data from *try and* pseudocoordination demonstrate the previously unsuspected need for explicit reference to morphological inflection in syntactic derivations. Put simply, the Inflectional Parallelism Requirement increases grammatical complexity. Modern English has a (complexity-increasing) principle that older stages of the language lacked.

10.3.3 *Pseudocoordination and complexity crosslinguistically*

Pseudocoordination, in one form or another, is widespread crosslinguistically. The previous section established that English and colloquial Faroese necessarily are more complex due to an inflectional parallelism requirement, but are other languages with pseudocoordination also therefore more complex? The available data in general do not permit us to give a firm answer to this question. While the *go*-type is especially common, the *try*-type with the particular requirement of inflectional parallelism is not documented elsewhere. Pseudocoordination is found throughout the Indo-European family (cf. Ross 2013a; Coseriu 1966) and Semitic as well (cf. Lillas-Schuil 2006: 80). Beyond these families, documentation of the phenomenon is limited, but even so examples (frequently considered 'serial verbs' or simply mentioned as anomalous coordinations) can be found in at least Oceanic (Manam: Lichtenberk 1983: 544–548; Lenakel: Lynch 1983), Khoisan (Sandawe: Eaton 2003; Khwe: Kilian-Hatz 2003: 203 '|òe'), Atlantic Niger-Congo (Supyire: Carlson 1994, ch. 8; historically in Akan: cf. Lord 1989: 199), Finno-Ugric (cf. Coseriu 1966), Caucasian (subordinator etymology in Mingrelian: cf. Harris and Campbell 1995: 290), and Korean (Kwon 2004), and probably many more languages. Although none of these languages are known to restrict inflection as in the English *try*-type, we might wonder whether any are more complex because of pseudocoordination. One candidate is Spanish, where pseudocoordinative *ir* 'go' has been analyzed as a "topic auxiliary" (cf. Arnaiz and Camacho 1999). Whether this and other examples actually do increase complexity will be left to future research.

Modern Greek is a language that has *try* in pseudocoordination, but with English *go*-type syntax without the bare form condition and instead allowing parallel inflection, as in (45). Therefore, like the English go-type, the Greek construction entails the completion of the action described by the second verb.

(45) Tí prospathṓ kai kánō? (Ross 2013a: 67)
what try.1s and do.1s
'What do I try and do?'

South African English also appears to lack the bare form condition, and, for that matter, any parallel morphology requirement. Examples can be found that extend the *try*-type of pseudocoordination to non-bare first verbs, so that the construction's distribution is essentially equivalent to that of *to*-infinitives, as shown in (46–47). Although these examples are rare and may not be grammatical for most speakers (cf. Ross 2013b: 125), they do present evidence for the non-requirement of parallel morphology (assuming they are not mere speech errors).

(46) Noeleen tries and find answers and solutions.
[from http://www.tvsa.co.za/, retrieved May 2013]

(47) It's a comical battle every night trying and get him to sleep in his own bed.
[from http://www.african-boerboel.co.za/, retrieved May 2013]

Therefore, Standard English is subject to a condition that is not shared by Modern Greek and South African English, which—everything else held equal—makes Standard English more complex. Alternatively, one could make the argument that therefore UG (and thereby *all* languages) must be more complex, even if there is only evidence in a minority of languages. The next step in the research program will be to consider more phenomena in more languages, working with (to the greatest extent possible) exhaustive descriptions of languages. In time, this methodology should be able to reveal languages that differ in complexity, or the lack thereof.

10.4 Summary and future prospects

This chapter has taken one construction, namely *try and* pseudocoordination in English, and shown how it contributes to linguistic complexity. I hope to have made a strong case for the importance of exhaustive descriptions in measuring linguistic complexity. We may at this point wonder whether it is a realistic goal to compare languages in terms of their relative complexity. Consider the implications if this goal were realizable and all languages turned out to be equally complex. We would naturally then go on to propose a mechanism, based either in UG or in language processing, to explain equi-complexity. On the other hand, if languages turned out to differ in complexity, we might still inquire about whether there are limits governing the degree of complexity differences possible. These are questions for future research.

11

Cross-linguistic comparison of complexity measures in phonological systems

STEVEN MORAN AND DAMIÁN BLASI

11.1 Introduction

An assumed truism in linguistics is that if the structure of a language simplifies in one place, it is likely to complicate in another (e.g. Martinet 1955; Hockett 1955; Aitchison 2000).[1] Although the complexity of subsystems may vary within a given language, another assumption is that these differences balance out cross-linguistically, so that all languages tend to be equally complex (e.g. Hockett 1958; Akmajian *et al.* 1979; Crystal 1987; McMahon 1994; Dixon 1997).[2] This notion is furthered by the long-held view that linguistic structures are not affected by geographic or societal factors, with vocabulary being an exception, e.g. Sapir 1912.

Recently, these assumptions about complexity have been challenged (e.g. McWhorter 2001; Kusters 2003; Dahl 2004, 2008; Hawkins 2004; Shosted 2006; Miestamo *et al.* 2008; Givón 2009; Sampson *et al.* 2009; Sinnemäki 2011), and have included findings of correlations between complexity and geographic or sociocultural settings (e.g. Perkins 1992; McWhorter 2001; Kusters 2003; Trudgill 2004; Hay and Bauer 2007; McWhorter 2007; Nichols 2009; Lupyan and Dale 2010; Sinnemäki 2011). These studies are controversial, however, because it is difficult for linguists to come to a consensus on an objective measure for complexity (e.g. Ansaldo and Nordhoff 2009; see also Sinnemäki 2011) and because of the socially sensitive issue of claiming that one language is more complex than another. Nonetheless, at least two common

[1] The idea that languages compensate for complexity in different subsystems has been termed, among other things, the *compensation hypothesis* and the *negative correlation hypothesis* (Shosted 2006).

[2] The idea that all languages are equally complex has also been given several labels, e.g. the *equi-complexity hypothesis* (Sinnemäki 2011), the *invariance of language complexity* (Sampson 2009b), or simply 'ALEC' (Deutscher 2009).

trends for measuring complexity have emerged: measuring relative complexity (the cost or difficulty of using a system) and absolute complexity (the number of parts of a linguistic subsystem).

Assessing the complexity of languages is a non-trivial task and to date there has been no clear methodology to do so at a cross-linguistic aggregate level. Researchers have concentrated on specific aspects of languages, mainly divided along classic distinctions such as syntax, morphology and phonology. In this chapter, we investigate the complexity of phonological systems using a large sample of languages. Under the rubric of absolute measures, researchers have formulated many claims about phonological complexity in terms of univariate or bivariate analyses, i.e. in terms of the shape of statistical distributions or the directions of correlations between two variables (e.g. Hockett 1955; Maddieson 1980, 1984, 2006, 2007; and Justeson and Stephens 1984). The former allows languages to be ranked in regard to a given complexity metric and the latter provides a method for testing compensatory relations between linguistic phenomena. What we find is that claims about distributions and correlations in the literature about phonological systems complexity, seem to be, at best case, dubious.

In our investigation, we first calculate univariate measures that have long been used by linguists, such as the total number of segments in languages.[3] We also apply a novel measure by reducing segment inventories into the minimal set of distinctive features needed to describe their phonemes.[4] The resulting distributions of these measures fail to exhibit any obvious parametric shape and the correlations that serve to defend positions about linguistic complexity do not appear, particularly when we look at the level of language family stock.

Next we look for correlations between different variables, both linguistic and non-linguistic, that have received attention in the literature, including consonants-vowels, obstruents-latitude, mean word length-segments and population-segments. What we find is that, looking at the broadest level (using all languages without any filters) we obtain interesting and significant but weak patterns.[5] However, when we zoom in on the language family stock-specific levels, these correlations turn out to be not significant. This lack of significance may result from either too few data points or may point to 'true' insignificance. The first might be true for small families, but it is unlikely to be the case for large families, such as Nilo-Saharan, Afro-Asiatic, Niger-Congo, Trans-New Guinea, Austronesian and Indo-European. Thus when

[3] More recently, the number of segments has been used as a proxy for a language's 'phonemic diversity' and has been correlated to factors such as increases in population size (Nichols 2009) and serial founder effects (Atkinson 2011). However, see the rebuttals in *Linguistic Typology* 15(2) and in Cysouw *et al.* (2012).

[4] Acquiring a phonology is a process of acquiring contrasts rather than inventories; thus changes in phoneme inventories are better understood in terms of contrastive features and phonological contexts (Kabak 2004).

[5] In this chapter the significance level was taken as $\alpha=0.05$.

controlling for language family stock, there is only fragmented evidence pointing to universal correlations and there are no general trends.

This chapter is outlined as follows. In section 11.2 we provide an overview of how complexity has been defined in the linguistics literature and specifically how absolute measures have been applied to phonological systems. In section 11.3 we describe the datasets that we have compiled together for our investigation. In sections 11.4 and 11.5 we present and discuss our results. In section 11.6 we put forward our conclusions.

11.2 Defining complexity

11.2.1 *Linguistic complexity*

There are two types of approaches for measuring linguistic complexity, which Miestamo (2006a) calls 'relative' and 'absolute'.[6] Relative approaches involve quantifying the cost or difficulty involved in processing or using language, whether as a speaker, listener, L1 acquirer or L2 learner. For example, a measure of relative complexity is proposed by Kusters (2003: 6), who defines complexity as the amount of effort a second language learner, or language community 'outsider', makes to become acquainted with a target language that he or she is unfamiliar with. Specifically, Kusters measures the effect of sociolinguistic changes on the complexity of verbal morphology by contrasting language learning effort of outsiders of four typologically different language families. One problem with relative complexity measures for typological research is that, as Kusters points out, they invite the question: 'What is complex to whom?' Answering this question involves identifying different roles (speaker or listener), different situations (acquirer or learner) and quantifying the numerous factors involved in producing, perceiving, acquiring or learning a language.

Figure 11.1 illustrates, with just four languages, how every spoken language is in a complexity relation with every other language from a language learning standpoint. For example, the native speaker of German learning one of the Chinese languages will have a more difficult time than he or she might have learning English or French. This is due to the genealogical relatedness of German and English (they share many cognates and grammatical features) and the areal proximity of German and French (sustained language contact has led to shared vocabulary). Contrast this situation with the Chinese speaker learning English, French or German. Not only are these three languages typologically very different from Chinese, but the Chinese speaker would have to learn from scratch essentially every word. Kusters (2003) attempts to

[6] Kolmogorov complexity and its variants have also been invoked as an ideal measure for complexity in some of the linguistic complexity literature (e.g. Juola 1998; Benedetto *et al.* 2002; Dahl 2009), but in practical terms, with its many problems of interpretation and implementation, "it is quite useless as a measure of the complexity of natural systems" (Shalizi 2006: 53). Furthermore, although a full argument against Kolmogorov complexity is beyond the scope of this chapter, let it suffice to say that it is not computable, meaning that one cannot devise a computer program to determine its value.

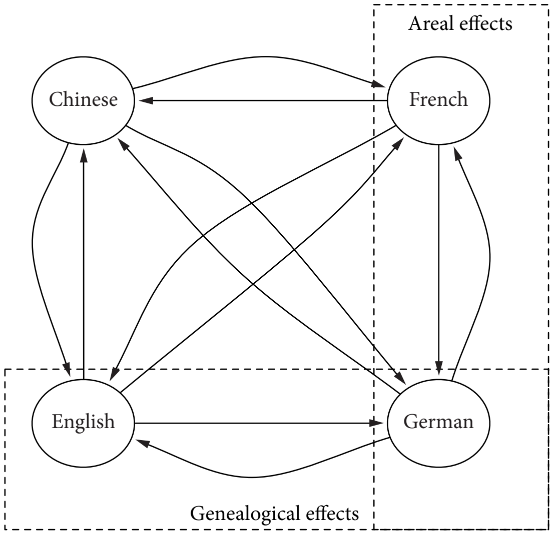


FIGURE 11.1 Relations and effects on L2 learner effort.

control for this effect by studying language outsiders like our hypothesized Chinese speaker. Nevertheless, for broad typological studies, measuring the effort of L2 learners is not currently feasible because every pair of languages will have its own unique bidirectional relationship that quantifies the relative complexity of the effort needed to learn the other language, complicated by factors of genealogical relatedness, areal proximity and elements of sheer coincidence, e.g. two languages that share similar sounds but which are not genealogically related or geographically close to each other.

Relative complexity metrics are exemplified by psycholinguistic studies that measure the cost and difficulty to language users in terms of processing or learning. The more difficult or costly it is to learn a linguistic phenomenon, then the more complex it is considered. It is hard, however, to define or quantify 'difficulty' of adult language learning in terms of metrics from psycholinguistic studies (McWhorter 2001; Trudgill 2001). For example, there are different burdens placed on the speaker and listener that require different metrics to capture articulatory- and perceptually-based effort as experienced by the speaker or listener. Therefore a phenomenon that is difficult for one group of language learners may actually be lessening the processing burden of another group. Miestamo (2009: 81) gives an example of discontinuous negation, as in French *Je ne chante pas*, which illustrates an extra production burden on the speaker, but eases the comprehension task of the listener. Another example is given by Kusters (2008), who argues that redundant agreement is difficult for L2 learners and speakers, but lessens the burden on L1 learners and hearers. Nevertheless, there is simply not enough data, well-defined methods and psycholinguistic research on all relevant aspects and phenomena associated with different relative complexity measures (Kusters 2008; Miestamo 2009). Moreover, one should also consider non-linguistic factors; Kusters (2003) argues that differences in complexity for outsiders are related to the social and cultural history of their own speech communities.

Given the problems involved in undertaking language acquisition studies or psycholinguistic experiments and quantifying and comparing relative metrics of complexity, Miestamo (2006a, 2009) suggests that cross-linguistic studies of complexity should instead define objective measures in absolute terms. Approaches to absolute complexity measure the number of parts of a linguistic system by some formal definition, often based on information theoretic approaches (e.g. McWhorter 2001; Dahl 2004; Miestamo 2006a, 2008, among others). Thus complexity is not seen as a metric of the difficulty or cognitive load of a speaker, listener or learner, but as a property of a linguistic (sub)system. In the following section we describe absolute measures that have been proposed to capture the complexity of phonological systems.

11.2.2 *The complexity of phonological systems*

Over the years, various absolute complexity measures have been used to describe and compare phonological systems. In this section we provide a historic overview of these measures, focusing on univariate and bivariate measures of complexity and compensatory relations in phonological systems. Univariate measures count the number of pieces in phonological systems, such as the number of contrastive sounds in a language, which can then be compared across languages. Bivariate measures use univariate counts and test for the presence or absence of compensatory relations between different phonological subsystems cross-linguistically; for example, do all languages with a great number of consonants have statistically fewer vowels? We also discuss the methodological considerations of these measures and show that the results of the same complexity metric have differed depending on the size and typological coverage of the language sample.

The first step in testing for complexity in phonological systems is to gather a cross-linguistic sample of language data. Once such a sample has been gathered, researchers usually turn their attention to the distribution of variables in the data, such as total segment counts. These discrete variable counts are identical to histograms. Especially interesting are the so-called 'parametric' distributions, which can be described with a finite (and typically small) number of parameters such as mean and variance. The parametric shape of distributions is of uttermost importance because it allows researchers to make inferences about the diachronic processes that produce them and the constraints that these processes are subject to (Maslova 2000; Cysouw 2007).

Segment inventory size is perhaps the oldest measure used for comparative surveys of cross-linguistic phonological patterns (e.g. Trubetzkoy 1939; Hockett 1955; Greenberg *et al.* 1978; Maddieson 1984). However, it is questionable whether overall phonological complexity can be captured by simply counting the total number of sounds in languages. Therefore researchers have divided segment inventories into different subsystems, such as consonant, vowel, and tone inventories. Each inventory can be further divided into different categorial distinctions, e.g. consonants by type (obstruent, glides, etc.), vowels by quality and non-quality distinctions, and tonal systems by categories such as simple and complex. Segment inventories can also be split into articulation classes (e.g. into sonorants and obstruents) or along phonetic dimensions (e.g. vowels by quality). These finer grained distinctions allow researchers to investigate phonological systems in more detail, but they also allow them to coarsen segment type distinctions into classes of sounds based on their distinctive features.

Apart from distributions, the different aspects of phonological systems are used for investigating hypothesized compensatory relations. One area that has been repeatedly investigated is whether there is a correlation between the size of consonant and vowel inventories. Using a sample of 68 languages, Hockett (1955: 138–140) was the first to investigate the 'balance' between the number of consonants and vowels. He found a correlation between the number of vowel phonemes and the total number of segmental phonemes. Maddieson (1980) obtained comparable results. With a genealogically stratified sample of 50 languages, Justeson and Stephens (1984) showed that the size of consonant and vowel inventories are statistically independent, concluding that there is no correlation between them. Using an even larger dataset with 317 languages, Maddieson (1984: 9) found that vowel ratio is inversely correlated with the number of consonants in an inventory and that the total number of vowels is positively correlated with the consonant total. In contrast, with an even larger dataset of over 600 languages, Maddieson (2007) more recently shows that consonant and vowel inventory sizes show no correlation with each other. And Moran (2012: ch. 5), using a dataset of around 1000 languages, calculates consonant and vowel ratios and shows that for each increase in roughly 13 or 14 consonants, there is an increase in one vowel. Although the p-value suggests that the finding is robust (slope = 0.0738, $R^2 = 0.0143$, $p < .0001$), the correlation is weak. Thus whether consonant and vowel inventories correlate in some way is still an issue of debate. This line of inquiry also shows how the results of one research question may change based on the sample of languages under investigation.

The consonant–vowel correlation has been actively investigated because it has long been suggested that languages compensate for complexity in one subsystem by simplifying another to achieve a roughly constant overall level of complexity. This 'compensation hypothesis' holds that a simplification or complication in one area of an inventory will be counterbalanced by the opposite somewhere else (Martinet 1955). These compensatory relations have been speculated as the outcome of historical processes that drive language change and they have also been investigated in light of languages being equally complex overall with regard to their communicative capacity (cf. Pellegrino *et al.* 2011). Maddieson (2006, 2007) investigates correlation patterns between different phonological subsystems in an attempt to shed light on compensatory relations in the phonological system, noting that regardless of whether the driving force is historical or communicative (or if it exists at all), if compensatory relationships hold they should present themselves in a survey of phonological properties of languages.

Table 11.1 illustrates Maddieson's (2006) findings from five phonological system variables taken to represent different metrics of complexity, which were then compared pairwise to see if they correlate positively (both measures increase in complexity), negatively (greater complexity in one measure; less in the other) or if they do not correlate at all. The variables for consonants, vowels, and vowel qualities are numerical, i.e. their values are represented by the total number of each type in the language under investigation. The variables for syllable complexity and tone complexity are categorial; they are based on the range of cross-linguistic types, which are then split into categories as defined by Maddieson (2006, 2007). Syllable complexity is based on the maximally complex syllable type found in the language and tone complexity is defined by the language's number and types of tones.

TABLE 11.1. Direction of correlation between complexity measures (Maddieson 2006: 106)

	Syllable complexity	Consonants	Vowel qualities	All vowels
Consonants	positive			
Vowel qualities	uncorrelated	uncorrelated		
All vowels	uncorrelated	uncorrelated	positive	
Tone complexity	negative	positive	positive	positive

By comparing different variables, Maddieson (2006, 2007) shows that the compensatory relationship in a sample of phonological systems holds only in the tone and syllable categories. He reports a positive correlation between consonant inventories and syllable complexity, tonal system complexity and consonants, and tonal systems and vowels (for both quality and full systems).[7] Maddieson finds no correlation for the following variable pairs: the number of vowels and syllable complexity, or the size of the inventories of vowels and consonants in languages in the sample. These findings are in line with earlier findings, where Maddieson (1984: 21) examines

[7] As Maddieson points out, there is an obvious positive correlation between vowel quality-only distinctions and full vowel systems.

suprasegmentals (tone and stress) in a set of 317 languages and reports that the "overall tendency appears once again to be more that complexity of different kinds goes hand in hand, rather than for complexity of one sort to be balanced by simplicity elsewhere."

The compensation hypothesis has also been tested outside of phonology. Nettle (1995) looks for a correlation between segment inventory size and word length, testing the assumption that as segment inventory size increases, the average word length in the language decreases. This relation might be thought to exemplify the communicative efficiency of a language, since a language with fewer sounds might need longer words to articulate the same number of concepts as a language with more sounds, because the latter has more combinatory possibilities for creating words. For his study, Nettle chose 10 languages and extracted their phonemic inventories from Campbell (1991). For each language Nettle determined an average word length figure by choosing 50 headwords at random from 10 representative dictionaries. Then Nettle compared whether the mean word length is related to segment inventory size and his findings show that correlation holds in the small data sample.

In another and more recent study testing for compensation between grammar and the phonological system, Shosted (2006) tests for compensatory patterning between verbal inflectional markers and potential syllables in a diverse sample of 32 languages. His approach is innovative, not only because he tests for a negative correlation across linguistic subsystems, but also because he develops an algorithm to calculate the combinatory possibilities of all syllable types in each phonological inventory by taking into account phonemic contrasts, syllable types and phonotactic constraints.[8] Shosted uses this calculation as a measure of phonological complexity, reasoning that there must be a discrete maximum syllable count, because elements of a phonological inventory are discrete. Thus Shosted addresses criticism of the approach taken by McWhorter (2001), who does not take syllabic structure into account when evaluating phonological systems complexity, and Shosted provides an empirical approach that can be retested and expanded. He finds no evidence for a negative correlation.

A major problem with many of the studies mentioned so far is their relatively small language samples. As pointed out above, whether the number of consonants and vowels are found to be positively or negatively correlated (or correlated at all) has differed depending on the study and the language sample. Clearly larger sample sizes are more desirable because they are more likely to contain a greater range of variability across languages, including marked and rare phenomena (the problem of course is gathering, analyzing and compiling such resources).

McWhorter (2001: 135) discusses complexity of phoneme inventories and defines their complexity through markedness of segment types, hypothesizing that a segment

[8] Maddieson (2009: 94) points out some problems with Shosted's method.

inventory is more complex than another if it has more marked segments. Markedness is calculated by a segment's cross-linguistic distribution, so marked segments are considered those that are less frequent in the world's languages. Again, a problem with this metric of complexity is sample size. For example, in the UCLA Phonological Segment Inventory Database (UPSID) there are 920 segment types that appear in 451 languages (Maddieson 1984; Maddieson and Precoda 1990). Of those 920 segments, 427 (about 46.5%) occur in one and only one language and therefore each is very rare in the sample. In the Phonetic Information Base and Lexicon (PHOIBLE) dataset of 1089 segment inventories, the number of total segment types occurring cross-linguistically increases to 1780. Of those 1780 segments, 909 (51%) are one-off occurrences. Thus in large samples of segment inventories created so far, nearly half of all segment types reportedly occur in just one language. Any study of phonological complexity that uses a small sample of languages is immediately suspected of lacking cross-linguistic representativity. The amount by which languages differ typologically, even within one subsystem, is quite astounding.

An additional methodological concern is the issue of how granular metrics of complexity should be when applied to phonological systems. As has long been noted in the phonological universals literature, phonemes are 'not fruitful' universals (Hockett 1963: 24), precisely because the analysis of allophones into a set of contrastive phonemes depends on several factors, such as the linguist's theoretical training and his or her ability to perceive sounds in the language being documented and described.[9] In evaluating Trudgill's (2004) thesis that phoneme inventory size correlates with social factors, Kabak (2004) argues that acquiring a phonology is not a process of acquiring an inventory, but of acquiring contrasts. Kusters and Muysken (2001) note that the phonological complexity approach taken by McWhorter (2001) does not take into account distinctive feature contrasts, but merely raw phoneme counts. Clements (2009: 29) provides an example of why simple segment counts do not inherently capture complexity by illustrating two historical stages of Zulu, which we replicate in Table 11.2.

Stage 1 reflects Zulu nearly a century ago and stage 2 shows a more modern usage of Zulu plosives (excluding fricatives). Both stages have the same number of segments, but through a historical change, the two stops (and sole members of their series) /ɓ/ and /k/ have shifted into a single voiced series. Clements (2009: 60) notes the increased feature economy of stage 2, which is due to the elimination of a feature (probably [obstruent]). What is clear is that although both stages of Zulu have the same number of segments, stage 1 requires a greater number of features to encode the phonemic contrasts in the language.

[9] Moreover, phonologists simply do not all agree on how to do phonemic analysis, as is apparent when comparing two or more segment inventories of the same language variety as described by different linguists (Moran 2012: ch. 2).

TABLE 11.2. Diachronic development of Zulu consonants (Clements 2009: 29)

Stage 1			Stage 2		
p’	t’	k’	p’	t’	k’
p^h	t^h	k^h	p^h	t^h	k^h
b	d	g	p	t	k
		k	b		g
ɓ					

We know of no previous work on complexity that uses the number of distinctive features as a metric for complexity. McCloy *et al.* (2013) use a multidimensionality reduction algorithm to reduce segment inventories in the PHOIBLE dataset into the minimal set of distinctive features needed to describe the phonemic contrasts in each language. In this chapter we use this dataset as an additional source for evaluating phonological complexity.

To summarize, we have described the absolute measures that have been used to measure the complexity of phonological systems and we have given an overview of some of the ways in which these measures have been used to investigate the compensation hypothesis. We have also raised some of the methodological concerns regarding language samples and how complexity measures have been formulated and tested on both phonological systems and other areas of the grammar. In the next section we describe the large dataset that we have compiled to test metrics of complexity and in sections 11.4 and 11.5 we run statistical tests for absolute complexity measures and discuss our results.

11.3 Data sample

The goals of our study are to test the absolute complexity measures of phonological systems and to compare previous findings against the results from a much larger and broader sample of the world’s languages. To achieve these goals, we first compiled together several datasets from different typological databases to create one large data sample for testing, which covers different units of phonological systems. In this section we describe this dataset and its component parts.

Contrastive segment inventories are the most widely documented aspect of the phonological systems in the world’s languages. This is due to the fact that the investigation and analysis of contrastive segments, i.e. phonemes, is the starting point for language documentation and analysis. Thus a great number of segment inventories have been compiled into segment inventory databases of varying genealogical and geographical coverage over the last several decades, among which are: the Stanford Phonology Archive (SPA; Crothers *et al.* 1979) and the UCLA Phonological

Segment Inventory Database (UPSID; Maddieson 1984; Maddieson and Precoda 1990). Phonological typology databases of this sort have provided a rich resource for investigating the segment inventory complexity of languages, allowing for measures such as the number of segments in a language and the ratio of consonants versus vowels.

In this study, segment inventory data is taken from the PHOIBLE database, which is comprised of SPA and UPSID. PHOIBLE contains several hundred additional inventories, bringing the current total number of languages in the sample to 1200 (Moran 2012). Inventories in PHOIBLE are 'factual claims' attributed to one or more linguists who either documented a language and used linguistic theory to posit a set of contrastive sounds for that language or consulted various published resources and inferred a set of contrastive sounds based on other linguists' analyses. The first type of factual claim is what one finds in secondary resources such as grammars and phonological descriptions. The second type is exemplified by the inventories in UPSID, in which Maddieson sometimes reinterprets secondary sources' phonemic analyses by a consistent standard. Both claims are influenced by various factors, such as the linguist's training, language background and the linguistic theory used to describe, analyze and posit contrastive sounds in the language under investigation. Thus all measures of complexity using contrastive phonological inventories are at their core based on researchers' assumptions, decisions and skills. In our experience, no two descriptions of the same language by different linguists contain exactly the same description of contrastive segments, so their segment types and counts may differ. A segment inventory is essentially an abstraction over the set of contrastive segments used by a particular language variety's phonological system, as defined by the set of distinctive features employed by that language (Clements 2009: 19). Whereas a segment is an abstraction of an articulatory or auditory unit of speech production or perception, features represent the basic phonetic units of a segment as defined by a particular distinctive feature theory. Therefore, in our study we move beyond just segments and on to the level of distinctive features that are needed to describe segment inventories, which allows us to abstract away from individual researcher preferences by using features based on articulation and perception.

The feature system used in this chapter is built on Hayes (2009), which was expanded in Moran (2012) to gain greater typological coverage of the segments in the inventories in the PHOIBLE database. It contains a set of 37 hierarchically organized features. To calculate the number of features needed to encode the contrasts in each segment inventory, McCloy *et al.* (2013) implemented a multidimensionality reduction algorithm that reduces the feature vectors of all contrastive segments in an inventory into the minimal number of features needed to encode that inventory. These sets of features constitute a data point in the sample of languages that we use in this study.

Additional information in the form of syllable structure data is taken from Maddieson (2011b) and available through the World Atlas of Language Structures (WALS; Dryer and Haspelmath 2011) and added to our compiled dataset. The WALS data sample contains 486 languages and syllable canon types are divided into three categorial values: simple, moderately complex and complex. Simple syllable structure is found in 61 languages that only allow (C)V. Another 274 languages with moderately complex syllables have CVC or CCV structures. And complex syllable structures occur in 151 languages that are defined as permitting "freer combinations of two consonants in the position before a vowel, or which allow three or more consonants in this onset position, and/or two or more consonants in the position after the vowel" (Maddieson 2011b: ch. 12).

Other than data about phonological systems, we include information about wordlists, genealogy, geography and demography. Mean word length figures are derived from Swadesh wordlists collected and compiled for the ASJP database (version 15) (Wichmann *et al.* 2012). Language family stock data come from the Ethnologue 15 (Gordon 2005) via Multitree (LINGUIST List 2009). Geo-coordinates for languages are taken from WALS (Haspelmath *et al.* 2008). Population figures are from Ethnologue 16 (Lewis 2009).[10]

11.4 Results of absolute complexity measures

First we examine the total number of segments per language in the PHOIBLE data sample. The distribution of segment inventory sizes by languages is given in Figure 11.2.

At least two proposals for a parametric model underlying this distribution have been put forth: log-normal (Justeson and Stephens 1984) and normal (Nichols 2009). While the normal (or Gaussian) distribution requires little introduction, the log-normal distribution (whose logarithm transform is normal) is characterized by an asymmetric shape around its peak, with a marked right skew. In practical terms this means that, in contrast with the symmetric and concentrated normal distribution, log-normal distributions have a non-negligible number of points far right from their maxima. Fits for normal and log-normal distributions on our dataset are shown in Figure 11.3.

A standard statistical test shows that neither the data (Shapiro-Wilk test, DF = 1299, $W = 0.894$, $p < 10^{-6}$) nor its logarithm transform (Shapiro-Wilk test, DF = 1299, $W = 0.994$, $p < 2 \cdot 10^{-4}$) are normally distributed.[11] However, it is important to

[10] As this study goes to press, new versions of WALS and Ethnologue (Lewis *et al.* 2013) have recently been published.

[11] The results from all normality and log-normality tests conducted here hold as well when the three most extreme segment inventories are discarded. These are: !Xóõ (161), Ju|'hoan (141) and Khwe (119).

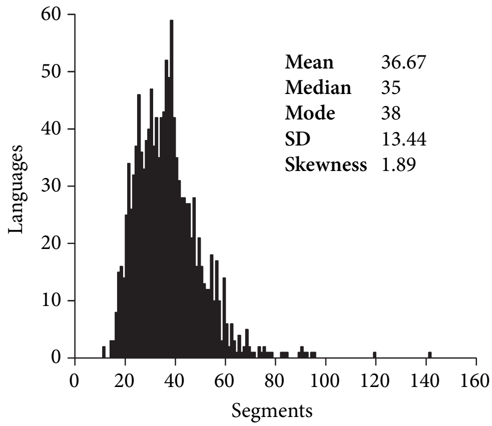


FIGURE 11.2 Distribution of segment inventory size in the PHOIBLE database.

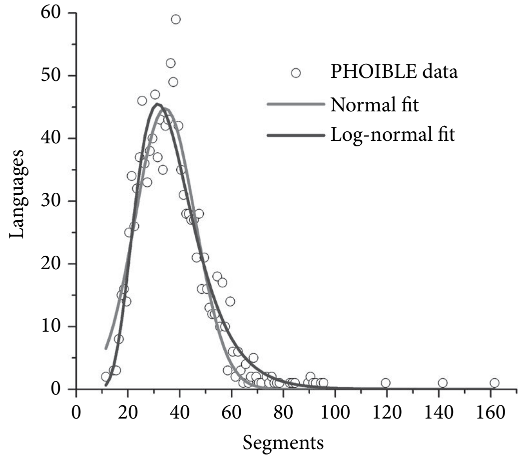


FIGURE 11.3 Normal and log-normal fit of the segments distribution in the PHOIBLE database.

remember that repertoire size can only be a natural number, whereas normal and log-normal distributions are defined in a continuous domain. Continuity is assumed as well in the normality test, and the violation of that condition may lead to erroneous results. For this reason we test how well the total number of segments can be adjusted by a negative binomial distribution, which is a general model for count data. Once more we find that the data fail to conform to any particular parametric model (Likelihood Ratio Test, $\chi^2 = 197.71$, $DF = 74$, $p < 10^{-6}$).

A closer inspection reveals a structure of sharp local maxima along the distribution, in contrast with the unimodal fit of both normal and log-normal distributions. The reason behind this fine-grained structure becomes clear when stock-specific

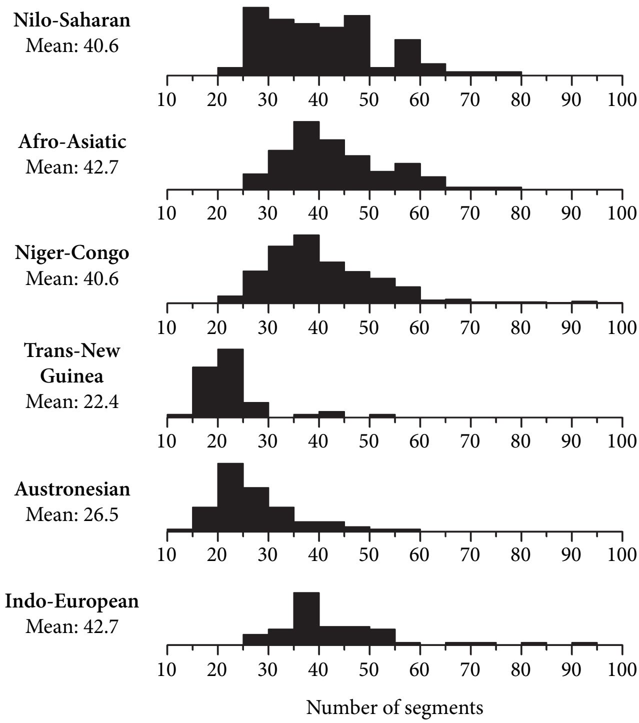


FIGURE 11.4 Distribution of segments for Nilo-Saharan, Afro-Asiatic, Niger-Congo, Trans-New Guinea, Austronesian, and Indo-European stocks.

distributions are analyzed, as is shown in Figure 11.4.[12,13] The differences in shape and average size of the inventories across families are remarkable, and a first look suggests that areal effects play an important role.

Most of what we have pointed out for segment distributions is also true for consonants and vowels when taken separately. The distributions for consonants and vowels in PHOIBLE can be seen in Figure 11.5. Although Justeson and Stephens (1984) propose a log-normal distribution for consonant and vowel inventories (based on a sample of 50 languages taken from Ruhlen 1975), the histograms show a much more complicated story.

Here again we find that, when studied individually, each language family has characteristic distributional features, as is apparent in Figure 11.6.

[12] We restricted our analysis to the six largest language families in PHOIBLE, which account for about 60% of the total languages.

[13] When no explicit scales are displayed, the areas of the distributions are modified for expositive purposes.

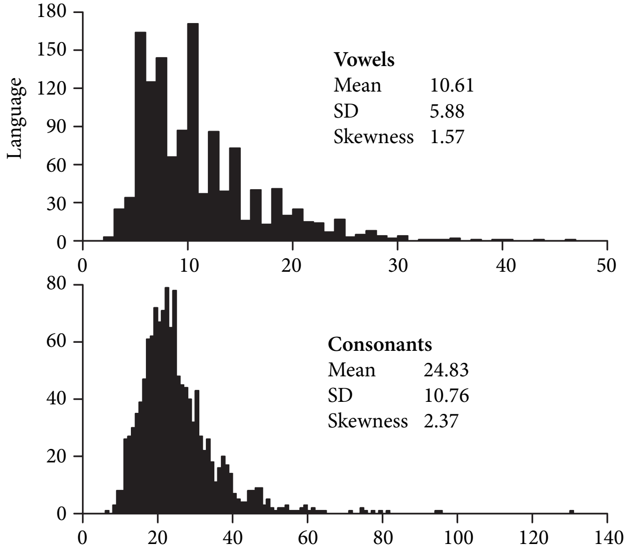


FIGURE 11.5 Distribution of consonants and vowels in the PHOIBLE database.

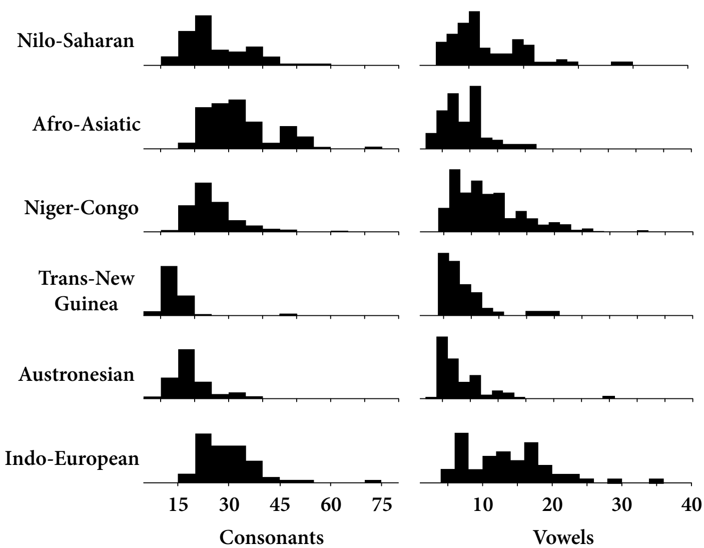


FIGURE 11.6 Distribution of consonants and vowels for Nilo-Saharan, Afro-Asiatic, Niger-Congo, Trans-New Guinea, Austronesian, and Indo-European stocks.

Together with Figure 11.4, Figure 11.6 illustrates why single representative sampling from a language family, as in Shosted (2006), is not appropriate. For instance one can see that the range for consonants (and total segment counts) for Nilo-Saharan, Afro-Asiatic and Niger-Congo is quite large: there is about a fourfold difference between the minimum and the maximum. However, Trans-New Guinea and Austronesian have a more concentrated distribution. A parametric model would force us to collapse many informative quantities into one: for instance, in peak distributions such as normal, Laplacian and Cauchy, the mean, mode and median all coincide. With no obvious parameterization for the families' distributions, then any single scalar (be it the mean, median, mode, any moment or any other statistic) will necessarily lack information about the diversity and the particularities of the distribution it comes from. Thus a single number is inadequate for summarizing the information of a whole stock.

So far we have presented observations about the data without considering the possible correlation between variables. First we look for any pattern between consonant and vowel counts. The total counts for consonant and vowel inventories are shown in Figure 11.7.

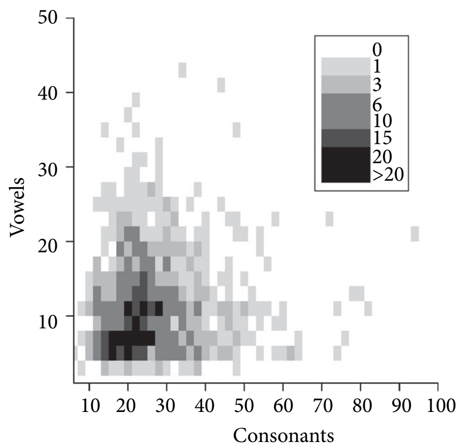


FIGURE 11.7 Vowel and consonant inventories in the PHOIBLE database.

The most salient trend is a mild, although significant, positive correlation between the number of consonants and the number of vowels (Spearman's ρ: 0.12, $p < 10^{-5}$). When we ask whether this effect holds within the largest language family stocks in the database, we find that only one individual family yields a statistically significant result. A linear correlation for Indo-European turned out to be significant and of moderate effect (Pearson's r: 0.31, p = 0.018).

These results should be considered carefully. It might be that the lack of significance is due to the small sample size, but given the fact that we have chosen particularly large language family stocks for our analysis, the reason is likely to lie elsewhere. Within most stocks under investigation, consonants and vowels do not exhibit any special pairing, in the sense that they show remarkably different numbers of consonants and vowels, from what one would expect in a random matching from distinct languages within the stock. Under these circumstances, the original correlation would be a mere by-product of the families' contrasting characteristics instead of denoting a causal connection between consonants and vowels.

Our next analysis concerns the set of features that constitute the set of contrastive segments in each inventory. We first consider the distribution of the number of minimum features required to describe a language's segmental repertoire, shown in Figure 11.8.

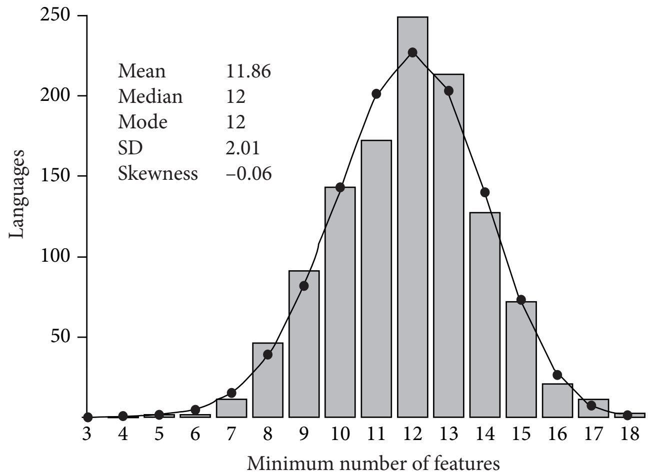


FIGURE 11.8 Distribution of the minimum number of features required to describe each segment inventory in the PHOIBLE database. Black dots correspond to the best fit for a binomial distribution.

The summary statistics of this distribution reveal a much more symmetric and balanced shape than that for segment inventory sizes. As we mentioned above, Justeson and Stephens (1984: 531) argue that both consonant and vowel inventories follow a log-normal distribution. They also claim that this distribution "receives straightforward linguistic interpretation in terms of distinctive features." If the number of features were connected with the number of segments through a logarithm, then from the log-normality of the segments we would be able to deduce the normality of the feature distribution. The distribution of the number of minimum features can be roughly approximated by a binomial distribution (which is the adequate parallel of the normal distribution in our case), although we must conclude that this model is still not satisfactory (Likelihood Ratio Test, $\chi^2 = 21.72$, $DF = 12$, $p = 0.04$).

However, the biggest issue in Justeson and Stephens's reasoning falls on the conjectured logarithmic relation between features and segments. Such a relation is true, for instance, in the case of a fully productive segment inventory, which makes use of all of the contrasts provided by its feature set. Thus if the feature set is size F, then the segment inventory will be of size $S = 2^F$. However, the relations between segments and features in this dataset exhibit a large degree of variability in the fraction of possible segments attested, as is shown in Figure 11.9.[14]

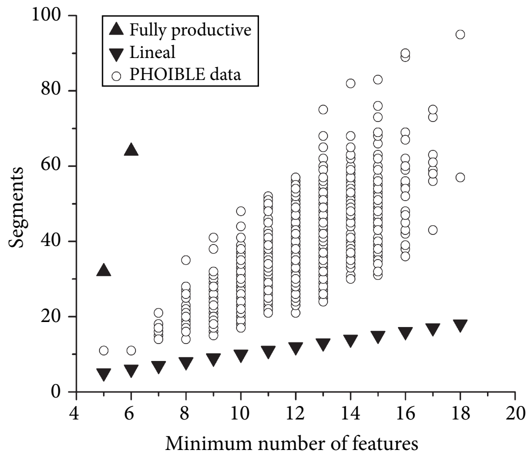


FIGURE 11.9 Segment inventory size vs. minimum number of features needed to encode each inventory.

Besides the form of the distribution of features, another relevant cross-linguistic property of features directly related to complexity is their rarity, i.e. how frequently they are attested in world's languages.

There is a group of six features that are highly represented cross-linguistically in our data sample, namely, dorsal (86%), high (83%), labial (83%), syllable (80%), continuant (80%) and voiced (79%). After that, the frequency of the remaining features decays in a manner that can be reasonably well approximated with an exponential fit (Adjusted $R^2 = 0.988$). This is shown in Figure 11.10.

Next we look at the relation between mean word length from wordlists in the ASJP database (measured in number of distinctive segments) and corresponding languages' segment inventories in the PHOIBLE database. Similarly to what was found by Nettle (1995), mean word length shrinks with larger segment inventories (Spearman's ρ: -0.309, $p < 10^{-5}$), but particularly with the number of vowels (Spearman's ρ: -0.311, $p < 10^{-5}$), instead of with the number of consonants (Spearman's ρ: -0.182, $p < 10^{-5}$). The effect is larger for small repertoires up to ten vowels, and

[14] As a comparison, we include points for a binary productive segment inventory (see text) and a linear one (where each feature grants a single segment in the repertoire).

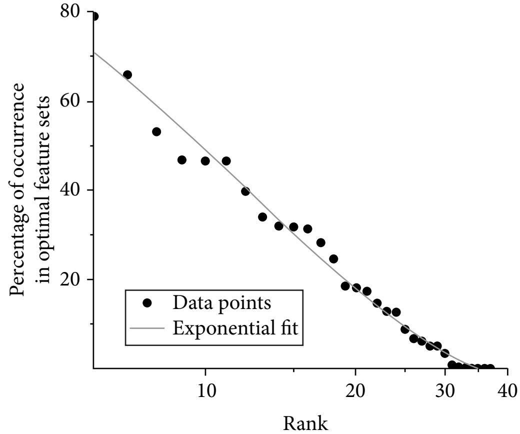


FIGURE 11.10 Ranking of features according to their cross-linguistic frequency and exponential fit (first 6 values not shown).

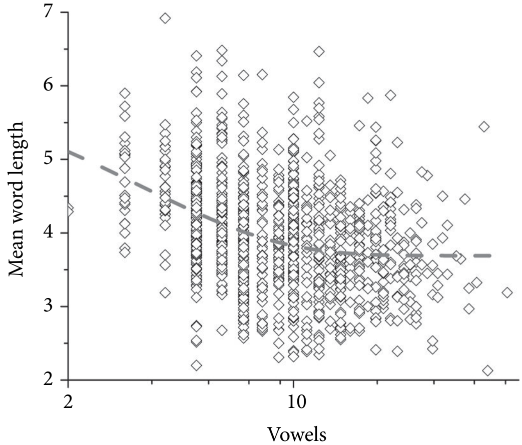


FIGURE 11.11 Mean word length (from ASJP) vs. vowel inventories (from PHOIBLE). The dotted gray curve is an exponential approximation of the moving average of data.

afterwards the mean word length stabilizes. The vowel-to-mean word length correlation is illustrated in Figure 11.11.

When language family stocks are considered in isolation, however, the vowel-mean word length correlations go beyond the threshold of significance in only two families: Indo-European (Spearman's ρ: -0.503, $p = 2 \cdot 10^{-4}$) and Niger-Congo (Spearman's ρ: -0.149, $p = 0.007$). These are illustrated in Figure 11.12. The other four stocks fail to reach significance, thus showing how widely the patterns differ.

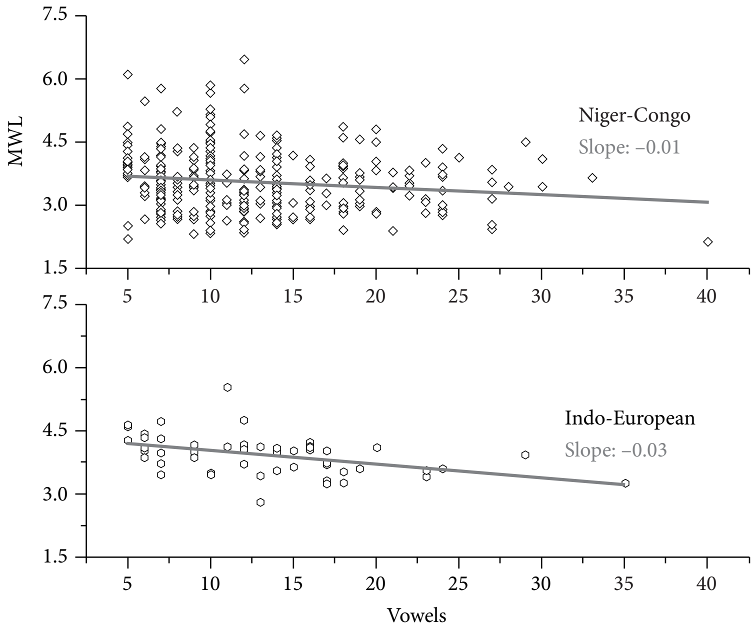


FIGURE 11.12 Mean word length versus number of vowels for Niger-Congo and Indo-European families. Straight lines correspond to the best linear fits.

Finally, we assess the influence of environmental and societal variables on some of the measures of absolute complexity. First we look for a correlation between latitude and various inventory sizes. Using a larger sample of languages than these previous studies, Maddieson (2011a: 32) finds a highly significant correlation between a phonological index (calculated as the sum of consonant inventory size and categorical syllable complexity values) and the distance from the central latitude of the world's temperate zones.[15]

Using the PHOIBLE dataset, we find a positive correlation between latitude and obstruents (Spearman's ρ: 0.451, $p < 10^{-5}$). Interestingly, vowels have a much weaker correlation (Spearman's ρ: 0.162, $p < 10^{-5}$) and sonorants do not even correlate significantly. This information is displayed in Figure 11.13, along with the best linear fits for illustrative purposes.

Latitude elicits the same tendency in syllable structure as well, as can be seen in Figure 11.14. Far north of the equator, syllabic structure is prominently complex, whereas close to it is mainly of moderate complexity, as defined in Maddieson (2011b) (Wilcoxon Rank Sum Test, all Holm-corrected $p < 10^{-6}$).

[15] The sample size is reported as between 450 and 650 languages.

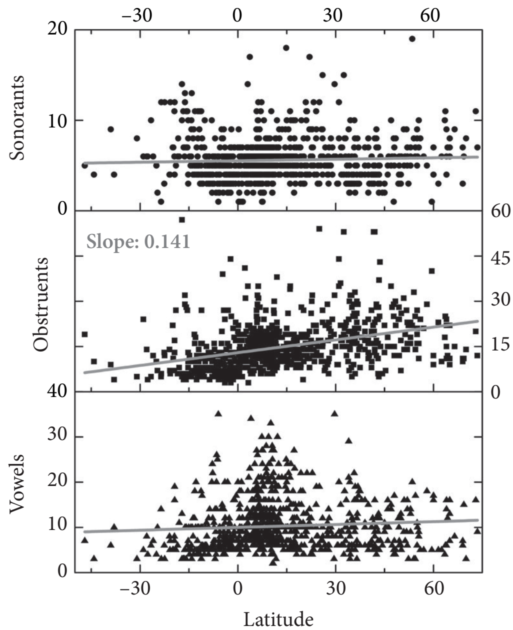


FIGURE 11.13 Sonorants, obstruents and vowels repertory size versus latitude. Best linear fits are displayed for reference. Slope values for vowels and sonorants are not displayed because they were not significantly different from zero.

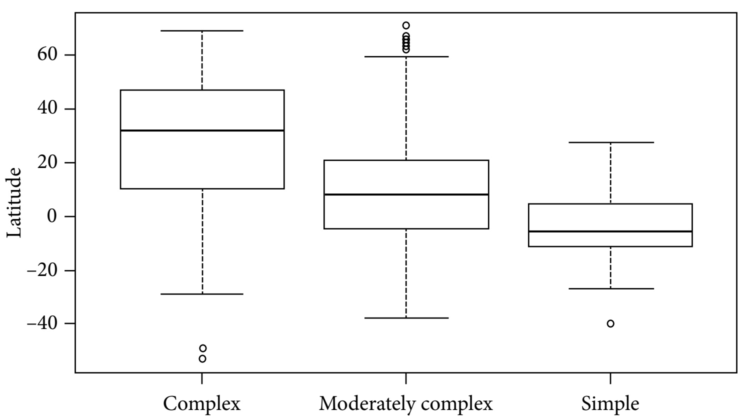


FIGURE 11.14 Complexity of syllable structure versus latitude.

The study of the effect of latitude on segments according to language family can be misleading, since in our database all but six stocks occupy a range larger than 30° latitude, namely Tupi-Guarani, Altaic, Afro-Asiatic, Austronesian, Indo-European, and Niger-Congo. Only the last four have enough data to let us infer a significant tendency. The picture that we obtain for the obstruents-latitude relation is mixed: Afro-Asiatic (Spearman's ρ: 0.32, $p = 0.027$) and Austronesian (Spearman's ρ: 0.233, $p = 0.043$) follow the overall trend, whereas for Indo-European (Spearman's ρ: -0.323, $p = .024$) and Niger-Congo (Spearman's ρ: -0.140 $p = 0.022$) larger latitude is often more related to a decrease in the number of obstruents.

Another important non-linguistic factor that purportedly shapes linguistic complexity, as measured by segment inventory size, is speaker population size (Haudricourt 1961; Trudgill 2004; Hay and Bauer 2007). Our results corroborate that the overall correlation between segments, consonants and vowels on the one hand, and population on the other, is positive, with total segments being the correlation with the largest statistic (Spearman's ρ: 0.331, $p < 10^{-5}$). This is shown in Figure 11.15.

Again, a family-specific analysis yields no significant correlations for most of the stocks except two, Afro-Asiatic (Spearman's ρ: 0.341, $p < 0.02$) and Austronesian (Spearman's ρ: 0.381, $p < 0.04$).[16]

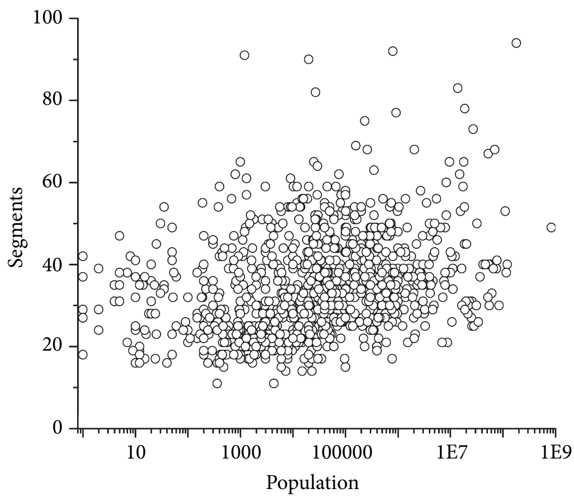


FIGURE 11.15 Segment repertoire size vs. population.

[16] In fact, a detailed analysis of this phenomenon was carried out by Moran *et al.* (2012), who introduced mixed effect models that can distinguish the linear contribution of the stock and population to the number of segments. Their findings show that, at least under a linear model, the contribution of population, independently of language stock, is not significant.

11.5 Discussion

In a nutshell, our results show that claims about either the parametric shape or the correlations among phonological variables appear as unsustained in the light of a simple statistical analysis of the largest phonological database to date. The distributions of the size of segment, consonant and vowel inventories fail to be classified by a standard statistical test as either normal or log-normal distributions (Justeson and Stephens 1984; Nichols 2009). Moreover, when these distributions are inspected in more detail, they reveal a rich structure of local peaks that have their origin in the superposition of the widely variable distributions of individual language family stocks.

We mentioned before that segment counts in the PHOIBLE database are 'factual claims', strongly influenced by the knowledge and the praxis of the linguist, and as such, segment counts are not the most reliable indices that can be used for the purpose of measuring complexity. Instead, we claim that counts of minimal features are a more reliable measure, as we discussed in section 11.3. This claim turned out to be backed by the analysis of the cross-linguistic frequency of occurrence of each of the 37 features from which minimal sets can be built. From a purely statistical point of view, the fact that these frequencies are well adjusted with an exponential function is relevant, because (in this case) it implies that most of the features will be present in many languages, and only a few will be cross-linguistically rare. This stands in contrast to the so-called 'heavy-tailed distributions' that usually have a large number of infrequent events. Conspicuously, segments constitute such a distribution: in both the UPSID and PHOIBLE databases, about half of all segment types are attested in one and only one language. Thus, feature counts are more conservative when it comes to measuring phonological complexity.

Taking together all the results concerning correlations between phonological variables, two commonalities appear. First, all pairwise correlations for the full dataset have a small effect, meaning that the statistical trends emerging from these data were far from revealing dramatic dependencies between variables. In other words, any individual variable does not give us sufficient information about another, e.g. knowing the number of consonants in a language does not give us power to predict the number of vowels in that language.

Second, and perhaps more importantly, for the largest stocks in the data sample used in this chapter, there was no uniform picture emerging from the pairwise correlations. Specifically, while some stocks do coincide in the direction of effect with the overall tendency, most stocks do not reveal any statistically significant trend. Under these circumstances, it is valid to call into question the universality of explanations linking different phonological subsystems. Presently, our data favor a historical (genealogical or areal) explanation behind most of the important correlations found. If there are cognitive and ecological factors molding the system of relations between phonological subsystems, they appear to be slower than the typical dynamics of language change.

11.6 Conclusions

In the last decade or so, the long-held assumed truism that all languages are equally complex has been challenged by many researchers. These researchers have explored issues of complexity from different perspectives and have proposed various metrics for quantifying complexity in and across linguistic subsystems. However, so far there are no conventionally agreed-upon metrics for quantifying or measuring complexity in the grammar and there has been no comprehensive quantitative method that has proven or disproven the notion of grammatical equality with regard to complexity.

In this chapter we have explored from a typological perspective several absolute measures used to quantify complexity of phonological systems. Despite differences of focus and methods, we have found that most of the research in this area relies on claims about statistical aspects of datasets, and particularly on the parameterization of distributions and the direction of correlations between pairs of variables. After compiling together datasets from some of the largest phonological databases available, we have shown that both strategies are problematic and to date no universal and salient trend can be established. First, distributions for the size of phonological systems are statistically different from the simple parametric models proposed in the literature, i.e. normal and log-normal. We have shown that the counts of minimal sets of features for segment inventories—first introduced in this work—are more robust than those of segments, and as such appear to be more accurate for the purpose of measuring phonological complexity. Second, pairwise analyses of phonological variables yielded significant, yet small correlations. If these correlations were propelled by universal laws of compensation or complexification, we should expect to see them within individual stocks as well—however, this is not the case.

As with most statistically-based research, our conclusions cannot rule out that a more detailed database or the choice of a particular set of variables could yield convincing universally-valid results regarding the structure or the development of phonological systems. However, it might be as well that our results are a consequence of languages being intrinsically flexible, adaptable structures that can allow a broad number of configurations and dynamics that do not fit well within a classification of universal types, cf. Evans and Levinson (2009), Bickel (in press).

12

The measurement of semantic complexity: how to get by if your language lacks generalized quantifiers

LISA MATTHEWSON

12.1 Introduction

Do languages vary in their semantic complexity? And if they do, does Universal Grammar place limits on that variation? These questions are not easy to answer. While there is a substantial literature investigating syntactic complexity in human language (see the other chapters in this volume, and references therein), semantic complexity has received much less attention, especially in the formal literature. We do not yet have an accepted definition of semantic complexity, nor do we know how to measure the semantic complexity of individual constructions, let alone entire languages. Empirically, we also have a problem: we know relatively little about how languages vary in their semantics in the first place.[1]

Given this state of the art, my goals in this chapter are relatively modest. My first aim is to illustrate how one goes about detecting semantic variation. I do this by means of a case study on quantification in St'át'imcets (Salish). Based on the St'át'imcets data, I argue (following Davis 2010, but contra Matthewson 1998) that languages differ significantly in the semantics of their quantifier-like elements. To be specific, some languages lack generalized quantifiers. The moral of this section is that semantic variation exists, but is not surface-visible or easy to detect.

[1] Thanks to St'át'imcets consultants Carl Alexander, Laura Thevarge, the late Beverley Frank, the late Gertrude Ned, and the late Rose Agnes Whitley. Thanks to Henry Davis and participants at the workshop on Formal Linguistics and the Measurement of Grammatical Complexity for helpful discussion, and to Fritz Newmeyer and Laurel Preston for comments on an earlier version of this chapter. This research was supported in part by SSHRC grants #410-2007-1046 and #410-2011-0431.

My second goal is to ascertain whether the attested variation in quantifier semantics involves a difference in semantic complexity. While it might seem obvious that a language which lacks generalized quantifiers is less complex than a language which possesses them, I will argue that this is not necessarily the case. A number of factors need to be taken into account, including formal complexity of the relevant lexical entries, paradigm complexity, and expressive complexity. I will conclude that the St'át'imcets quantificational system is in fact less complex than that of English, but not for the reasons one might have thought.

Finally, I present a brief language-wide comparison between English and St'át'imcets, assessing (as far as is possible with current knowledge and tools) the two languages' relative semantic complexity in a range of areas, including tense, aspect, evidentials and so on. Overall, the competition is pretty much a tie, with each language balancing paradigm complexity in some areas with simplicity in others. Is this balancing a coincidence? Could some languages be, overall, semantically more complex than others? While balancing makes functional sense, I argue that Universal Grammar has no means to prevent languages from differing in overall complexity. Universal Grammar places formal and substantive constraints on meaning, but not on semantic complexity per se. Universal Grammar does not even prevent differences in effability; thus, some languages are able to express meanings which other languages cannot (von Fintel and Matthewson 2008; Deal 2011). This claim is supported by the English/St'át'imcets comparison presented here, as the languages display at least one effability difference.

12.2 How do languages vary in their semantics?

Before we can ask whether languages vary in their semantic complexity, we need to ascertain how languages vary in their semantic properties, period. This is an area about which surprisingly little is known. Although cross-linguistic formal semantic research has been increasing rapidly since the 1990s, there is no accepted theory of what is universal and what varies in semantics. There is not even any active debate about what limits are placed on semantic variation. Instead, we have on the one hand isolated formal analyses of particular phenomena in different languages (often only a handful of languages), and on the other hand, large-scale typological research which relies on sources whose data were not gathered for the purpose of testing empirical hypotheses about semantics (see for example Dryer and Haspelmath 2011). As argued in Davis *et al.* (in press), the latter type of research usually fails to provide useful results for semanticists, as it offers only incomplete information about meaning. Davis *et al.* argue that what is needed is formal cross-linguistic research on as many languages as possible, testing concrete hypotheses, leading eventually to a formal semantic typology.

Exciting work is already going on in this vein, which will eventually lead us to an informed picture of universality and variation in semantics.[2] For now, we can say that there is restricted cross-linguistic variation in the inventory and lexical semantics of functional morphemes such as aspects, tenses, evidentials, modals, quantifiers, and determiners. The variation may be non-trivial, but it is not without limit or constraint. Semantic variation seems to be usually (but not always) non-parametric, and often consists of differences in which of a set of core building blocks of meaning are utilized, and/or in how finely languages lexically divide up the same semantic space (see von Fintel and Matthewson 2008 for discussion).

Let's turn to our case study on quantifiers.

12.3 Quantifiers

In this section I present a case study in semantic variation, that of quantifiers. Quantification is relatively well-studied from a formal, cross-linguistic perspective: see, for example, the collections in Bach *et al.* (1995), Matthewson (2008), and Keenan and Paperno (2012). Here I do not attempt to overview the state of the art, but concentrate on one specific empirical question, namely, whether some languages lack generalized quantifiers (GQs), and I look at just two languages, English and St'át'imcets. The question is of relevance to complexity because a language which lacks GQs may be less complex than a language which possesses them.

The answer to the empirical question (following Davis 2010) is yes, some languages do lack generalized quantifiers. A bonus moral of the discussion will be that one cannot establish the absence of a certain kind of quantifier by looking superficially, or by relying on translations or morphological make-up (cf. Everett 2005 on Pirahã). As for the complexity question, I will argue that a language which lacks GQs does *not* necessarily have a simpler quantification system than one which possesses them. In the case of St'át'imcets, however, there happens to be a lack of paradigm complexity in the quantificational system.

First, we need a definition of generalized quantifiers, so we know what to test for. A GQ is a noun phrase which enforces a particular relation between two sets, usually provided by a common noun and a VP. In (1a), for example, we see that *every girl* enforces a subset relation. *Every girl dances* asserts that the set of girls is a subset of those who dance. *Most girls* in (1b) enforces a majority relation. *Most girls dance* asserts that more girls dance than don't dance.

(1) a. $[[\textit{every girl dances}]] = 1$ iff $[[\textit{girl}]] \subseteq [[\textit{dance}]]$
    b. $[[\textit{most girls dance}]] = 1$ iff $|\, [[\textit{girl}]] \cap [[\textit{dance}]] \,| \geq |\, [[\textit{girl}]] - [[\textit{dance}]] \,|$

[2] Recent examples include Cover (2011), Deal (2011), Tonhauser (2011), McKenzie (2012).

Noun phrases of this type, which enforce a relation between two sets, are assumed not to be analyzable as denoting either individuals or predicates. Thus, a noun phrase of the form *every N* can neither be analyzed as being of type e, nor as being of type <e,t> (see e.g. Heim and Kratzer 1998). Instead, *every N* must be of type <<e,t,>,t>, denoting a function from sets/one-place predicates to truth values. We call noun phrases of type <<e,t>,t> 'generalized quantifiers' (see Barwise and Cooper 1981; Szabolcsi 2010; among many others).

When asking whether some languages lack GQs, what matters is whether the language has noun phrases which *must* be analyzed as GQs, not noun phrases which *can* be analyzed as GQs. Thus, we are interested in 'essentially quantificational' NPs, in the sense of Partee (1995). The reason for this is that any noun phrase *can* be analyzed as a GQ, including even proper names (Montague 1974). For example, the name *Henry* can be viewed as denoting a function of type <<e,t>,t>; the function returns the output 'true' for all sets of which Henry is a member, and false for all sets of which he is not a member.

How can we detect essentially quantificational noun phrases empirically? Two common diagnostic properties for GQs are proportional readings, and scopal ambiguities with respect to other scope-bearing operators. With regard to proportional readings, the English GQ *many A* has a reading where '*Many* A B' requires a large proportion of As to B. This reading is facilitated by a partitive structure, as in (2). If 25 out of 30 students passed the test, then (2) is fine. It is decidedly more questionable if 25 out of 100 did.

(2) Many of the students passed the test.

Many also has a cardinal reading—facilitated in (3) by a *there*-insertion structure—in which it suffices for there to be a large number of Ns. However, this reading does not require a quantificational analysis, as noted by Milsark (1974), Partee (1988), Szabolcsi (2010), among others.

(3) *Context: There are 40,000 students at UBC. Yesterday there was a protest rally and 2,000 students turned up.*
There were many students at the rally yesterday.

The second diagnostic for GQs is their giving rise to scopal ambiguities, as illustrated in (4).

(4) All the children built a raft.
Surface scope reading: For each child x, there is a raft y, and x built y. (different rafts)
Inverse scope reading: There is a raft y, and all the children built y. (only one raft)

As discussed by Davis *et al.* (in press), proportional readings and scopal ambiguity are both necessary conditions for an essentially quantificational noun phrase, and together, they constitute sufficient evidence. That is, a noun phrase which displays proportional readings and participates in scopal interactions must be analyzed as a GQ.[3] An initial hypothesis that a language possesses GQs is therefore falsifiable by determining that noun phrases in that language lack these properties. In the next subsection we attempt to find such evidence in St'át'imcets.

Let us turn now to evidence that St'át'imcets lacks GQs. The discussion relies in large part on the insights of Davis (2010). St'át'imcets (a.k.a. Lillooet) is a highly endangered Northern Interior Salish language spoken in the southwest interior of British Columbia. The methodology of data collection for the material presented here was one-on-one intensive fieldwork. Standard elicitation methodologies were used, including primarily a Felicity Judgment Task, in which speakers judge the acceptability of sentences in specific discourse contexts (Matthewson 2004, 2011).

St'át'imcets possesses several elements which appear quantifier-like, including *tákem* 'all', *zí7zeg'* 'each', *sáq'ulh* 'half', *cw7it* 'many', *k'wík'wena7* 'few' and *t'qwaw's* 'both'.[4] Matthewson (1998) argues that although these elements differ from English quantifiers in that they must appear inside a DP which also contains a determiner, the entire resulting phrase denotes a generalized quantifier. One piece of evidence for this claim is the presence of proportional readings. Consider (5), containing *cw7it* 'many':[5]

(5) úxwal' [i=*cw7ít*=a plísmen]
go.home [DET.PL=*many*=EXIS policeman]
'Many of the policemen went home.' (Matthewson 1998: 304)

(5) is rejected in a context where 25 out of 100 policeman go home, but accepted if 25 out of 30 policeman go home. Based on a range of similar data, Matthewson (1998) argues that *cw7it* (as well as *k'wík'wena7* 'few') has *only* a proportional reading when it appears inside DP. This appears to be evidence that the resulting noun phrases are essentially quantificational.

However, Davis (2010) shows that phrases containing elements like *cw7it* do *not* satisfy the other diagnostic for GQ-hood: they do not participate in scopal interactions.[6] He argues on the basis of this fact that St'át'imcets lacks generalized quantifiers.

[3] On its own, scopal behavior is not a sufficient condition, since elements other than GQs (e.g. negation, modals, intensional verbs) participate in scopal ambiguities. Matthewson (1998) assumed (following much prior literature) that proportional readings alone were a sufficient condition for generalized quantifierhood. We will see below that this is not correct.

[4] St'át'imcets data are presented in the orthography designed by Jan van Eijk and used by community members; see van Eijk and Williams (1981). The symbol *7* represents a glottal stop.

[5] Morpheme glosses not covered by the Leipzig Glossing Rules include DEIC = deictic; DEON = deontic; EMPH = emphatic; EPIS = epistemic; EXIS = assertion of existence; RED = redirective applicative; STAT = stative; TEMP = temporal.

[6] Matthewson (1999) showed that sentences containing one quantified phrase and one plain indefinite do not display scopal ambiguity. Davis takes the investigation a step further by looking at sentences containing *two* overtly quantified phrases.

Consider (6). Here, we have two apparently quantificational noun phrases, 'all the children' and 'half the books'.

(6) *Context: Four children are meant to read four books over the summer holidays.*
[*tákem* i=sk'úk'wmi7t=a] paqwal'ikst-mín-itas [*sáq'ulh* i=púkw=a]
[*all* DET.PL=child=EXIS] read-RED-3PL.ERG [*half* DET.PL=book=EXIS]
'All the children read half the books.'

Davis argues that neither of the two potential scopal readings exists for (6). The scenario in (7) is one in which (6) should be accepted under a wide scope reading for the subject: For each child x, x read a (different) half of the books. However, consultants reject (6) in this context.

(7)

A reads	B reads	C reads	D reads
books 1,2	books 2,3	books 3,4	books 1,4

(8) represents a context in which (6) should be accepted on an inverse scope reading: For a particular half of the books (1 and 2), each child read that half. However, consultants also reject (6) in the context in (8).

(8)

A reads	B reads	C reads	D reads
books 1,2,3	books 1,2,4	books 1,2,3,4	books 1,2

It turns out that (6) is accepted in all and only situations in which all the children read at least one book, and a *total* of two out of the four titles are read.[7] These contexts correspond to a *cumulative* reading. As discussed by Scha (1981), cumulative readings are scopeless readings, in which a predicate relates two sets which each have a certain total size. In our example, what is required is that all the four children participated in reading, and a total of half the four books were read (but exactly who read what is left vague). Scha (1981) shows that this reading cannot be captured using an ordinary GQ analysis. See Davis (2010) for data involving other combinations of quantifiers in St'át'imcets, showing in each case that the cumulative reading is the only available reading.

Davis's findings are quite significant: they mean that a St'át'imcets sentence containing two quantifier-like elements does not have the same semantics as its translation 'equivalent' in English. St'át'imcets really lacks the GQ readings of all such sentences. Notice that the absence is far from immediately obvious. Any sentence

[7] Support for the claim that it is the total number of books which matters comes from one consultant's comment on context (7): "No—they read all the books, so you couldn't say they read half the books" (Davis 2010).

with just one quantifier will seem similar to its English translation. For example, the St'át'imcets version of 'All the children ate fish' will, just like the English version, require every contextually salient child to have eaten fish. One can only detect the absence of GQ readings by means of the targeted fieldwork done by Davis.

Aside from the absence of GQs, St'át'imcets displays another absence in the area of quantification: there are no true generics. As mentioned earlier, Matthewson (1998) shows that quantifier-like elements in Salish always co-occur with determiners. Importantly, these determiners explicitly limit the domain of quantification by placing a deictic restriction on the individuals being quantified over. For example, (9a) contains the 'absent' plural determiner *nelh . . . a,* because the people being talked about are deceased at the time of speech, while (9b) contains the 'present' determiner *i . . . a* and picks out a particular group of present people.

(9) a. tqílh=wi7 tákem pináni7 *nelh*=ucwalmícw=a láku7
almost=EMPH all TEMP.DEIC *DIST.DET.PL*=person=EXIS DEIC
lil'wat7úl=a snek'wnúk'wa7-s
Mount.Currie=EXIS relatives-3POSS
'He was related to almost everyone in Mount Currie.' (Matthewson 2005: 393)

b. tqílh=wi7 tákem *i*=ucwalmícw=a láku7 lil'wat7úl=a
almost=EMPH ALL *DET.PL*=person=EXIS DEIC Mt.Currie=EXIS
snek'wnúk'wa7-s
relatives-3POSS
'He is related to almost everyone in Mount Currie.'

The explicit domain restriction which is obligatory with St'át'imcets quantification is incompatible with generic quantification, which by its very nature makes a claim about an entire class of individuals. Matthewson (1998) shows that attempts to elicit generic statements in St'át'imcets result in ordinary universal constructions, as in (10). She points out that the quantified phrase in such statements is identical in structure to quantified phrases involving a specific set of individuals, as in (11).

(10) *tákem* i=twéw'w'et=a ama-mín-itas k=wa píx-em'
all DET.PL=boys=EXIS good-APPL-3PL.ERG DET=IPFV hunt-INTR
'All boys love hunting.' (Matthewson 1998: 333)

(11) *tákem* i=cácl'ep=a twéw'w'et nás=tu7 píx-em'
all DET.PL=Fountain=EXIS boys go=then hunt-INTR
'All the boys from Fountain went hunting.' (Matthewson 1998: 333)

It appears that although St'át'imcets speakers can translate English generics into their language, the sentences they produce are not truly general. Support for this is given by (12). A speaker commented about this sentence that "there's a bunch of men there;

it doesn't pertain to all the men in the world." "However, when asked how she *would* refer to all the men in the world, [12] was the only way it could be done" (Matthewson 1998: 333). See also Gillon (2006) for similar evidence that the related language Sk̲wx̲wú7mesh lacks generics.

(12)	léxlex	s=Henry	[lhél=ki=*tákem*=a	sqáyqeycw]
	intelligent	NMLZ=Henry	[from=DET.PL=*all*=EXIS	men]
	'Henry is the most intelligent of all the men.'			(Matthewson 1998: 333)

The discussion in this section has shown that at least one language exists which lacks essentially quantificational noun phrases. We also saw that some languages lack generic readings. The question now is what the implications of these findings are for complexity. In the next section I lay out some criteria for measuring semantic complexity, and in section 12.5 I apply the criteria to the St'át'imcets quantificational system.

12.4 What is semantic complexity?

As mentioned above, there has been relatively little formal work on semantic complexity, and there is certainly no generally accepted definition. Semantic complexity is mainly discussed in the acquisition, psycholinguistic, or processing literature, and in these contexts it is usually not explicitly defined. Here I will lay out some possible ways in which we could measure semantic complexity.

Consider a fairly standard picture of what is included in the semantic competence of a native speaker of a language. The speaker must know at least (a) the lexical entries for all the morphemes in their vocabulary, (b) the mechanisms which are used to compose meanings together, in a way which interfaces with syntax, and (c) how to form pragmatic inferences (such as presuppositions and implicatures) and interpret context-dependent meaning (indexicality, vagueness, etc.). A speaker who knows all this will know, for any utterance in his or her language, its truth conditions, along with any pragmatic inferences it gives rise to in particular discourse contexts.

This picture points to a few different levels at which we could measure complexity. We could calculate the complexity of the formal representation of lexical entries or of the truth conditions of entire utterances. Another possibility would be to calculate the complexity of everything that goes into creating truth-conditional meaning (including composition mechanisms). We could also measure the paradigm complexity of particular areas of the grammar, by checking whether more or fewer semantic distinctions are expressed in the same semantic domain. We could include pragmatic inferences as part of complexity, and finally we could consider overall expressive complexity (whether some languages lack certain meanings altogether).

Let us first consider how to measure the complexity of formal representations of truth conditional content. There are two obvious ways in which this could be done: in terms of the length of the description (cf. Miestamo 2008, among many others), or in terms of semantic strength.[8] An example of the first approach comes from Gennari and Poeppel (2003), who assume that verbs denoting externally caused events (such as *break*) are more complex than verbs denoting internally caused events (such as *grow*). The reasoning is that "externally caused verbs, but not internally caused ones, conceptually require two participants" (Gennari and Poeppel 2003: B28) and hence their lexical entries are longer.

One problem with the length criterion is that there are often many different ways to write logically equivalent semantic formulas. For example, the two Russellian lexical entries for the English definite article in (13) are different lengths, but are equivalent.

(13) a. $[[\ \textit{the}\ ]] = \lambda f_{<et>} . \lambda g_{<et>} . \exists x\ [\forall y\ [f(y) \leftrightarrow x = y]\ \&\ g(x)]$
b. $[[\ \textit{the}\ ]] = \lambda f_{<et>} . \lambda g_{<et>} . \exists x\ [f(x)\ \&\ \forall y\ [f(y) \rightarrow x = y]\ \&\ g(x)]$

Another example is given in (14). These two possible denotations for *most,* which differ slightly in length but which are truth-conditionally equivalent, are discussed by Hackl (2009).

(14) a. $[[\ \textit{most}\ ]](A)(B) = 1$ iff $|A \cap B| > |A - B|$
b. $[[\ \textit{most}\ ]](A)(B) = 1$ iff $|A \cap B| > \frac{1}{2}\ |A|$

If the correct truth conditions can be obtained from two different formal analyses, we cannot measure the complexity of an element based on its formal properties until we have found a way to choose between the equivalent formal analyses.[9]

What about the 'semantic strength' criterion? Gennari and Poeppel (2003) also use this one. They assert that complex elements are those which have 'more entailed properties' than simpler ones. Thus, eventive verbs are more complex than statives, because "eventive verbs entail simpler conceptual units such as CAUSE, BECOME or CHANGE and resulting STATE, corresponding to the event's internal dynamics they denote, while stative verbs lack any such causal entailments" (Gennari and Poeppel 2003: B28). However, under an analysis in which eventive verbs contain operators like CAUSE or BECOME, the lexical entries for eventive verbs are also longer than those for stative verbs. In this case, Gennari and Poeppel's second criterion (number of entailments) can be reduced to the first (length).

The length criterion and the entailment criterion do not always give the same results, however. Consider numerals. The lexical entries for all numerals are the same length, yet (15a) entails (15b) and (15c), (15b) only entails (15c), and (15c) entails none

[8] An element A is stronger than an element B if a sentence containing A asymmetrically entails a parallel sentence containing B.

[9] Hackl (2009) uses processing evidence to distinguish between the two potential analyses in (14). He argues that only (14a) is correct for *most,* while (14b) represents the meaning of *more than half.*

of the others. It seems unlikely that we want to conclude that *four* is more complex than *three*. This suggests that semantic strength does not (alone) determine complexity.

(15) a. There are **four** cabbage trees in the park.
b. There are **three** cabbage trees in the park.
c. There are **two** cabbage trees in the park.

Another case in which length and semantic strength do not match up involves the English past tense compared with the St'át'imcets non-future tense (as analyzed by Matthewson 2006b). As seen in (16–17), the formulas contain the same number of symbols, and thus are equally complex according to the first criterion. However, a semantic strength criterion would classify the English past as more complex, since a sentence containing a dedicated past tense entails one containing a general non-future tense, but not vice-versa.

(16) $[[\ \text{PAST}_i\ ]]^{g,t}$ is only defined if $g(i) < t$. If defined, $[[\ \text{PAST}_i\ ]]^{g,t} = g(i)$.
(adapted from Kratzer 1998)

(17) $[[\ \text{NON-FUT}_i\ ]]^{g,t}$ is only defined if $g(i) \leq t$. If defined, $[[\ \text{NON-FUT}_i\ ]]^{g,t} = g(i)$.

Although semantic strength per se is probably not a good complexity measure, semantic strength often correlates with another measure, namely, paradigm complexity. Paradigm complexity embodies the idea that for the same semantic space S, a language which encodes sub-divisions within S is more complex than a language which does not (McWhorter 2001; Kusters 2003; Miestamo 2006b, 2008; Lindström 2008, among others). By this criterion, the English tense system is more complex than the St'át'imcets one, because there are more tenses in English than there are in St'át'imcets. Notice that a system which encodes more distinctions within a given domain will also tend to have elements with more specific (hence semantically stronger) lexical entries.

So far we have dealt only with truth-conditional content, and have ignored complexity of pragmatic inferences.[10] Some pragmatic phenomena, such as presuppositions and conventional implicatures, straightforwardly add formal complexity. For example, the Fregean analysis of the English definite article in (18) contains the underlined presuppositional material. This is the information which must be common knowledge in the discourse situation at the time of utterance.

(18) $[[\ \textit{the}\ ]] = \lambda f_{<e,t>}\ \underline{\text{and there is exactly one x such that f(x) = 1}}$. the unique y such that f(y) = 1 (Heim and Kratzer 1998: 75)

[10] I set complexity of composition mechanisms completely aside, as hardly any research has been done into whether languages can vary in their composition rules (see Chung and Ladusaw 2004 for one exception, and Gil 2005b for a more radical view of how languages can vary in composition strategies).

On the other hand, conversational implicatures do not add formal complexity, since they are not represented in the truth conditions.[11] Yet intuitively, an utterance for which an implicature is generated should count as more complex than an utterance without one. Some researchers are attempting to detect the added complexity of conversational implicatures using processing evidence (Bott and Noveck 2004; Breheny *et al.* 2006, among others). The idea is that conversational implicatures must be calculated, and this user complexity should have an effect on processing times.

However, research into the processing of pragmatic inferences is in its infancy, as pointed out by Sedivy (2007: 481). Sedivy writes (2007: 493): "We do not yet have a clear picture of the hearer's processing costs of computing implicatures; while some studies show a measurable processing cost, others suggest that Gricean inferencing can be used without detectable cost."

Another type of pragmatic 'work' which may add to complexity involves the inferences required by semantically underspecified or general elements. For example, we saw above that the St'át'imcets non-future tense is semantically weaker and paradigmatically simpler than the English past tense. However, precisely because the St'át'imcets non-future tense is more general, it may impose more demands on the hearer (who does not have the temporal reference narrowed down for them in the semantics). And indeed, semantically general elements can be shown to have a processing cost (Breedin *et al.* 1998).[12] Again, though, research is just beginning, and is heavily concentrated on English.

When thinking about the role of processing evidence, it is important to remember that merely measuring processing times (without having any formal analysis of the relevant constructions) would not lead to interesting results. This is particularly so since our different complexity measures can make different predictions for the same elements. A long (complex) description may correspond to a semantically strong (complex) truth-conditional contribution, which however requires little (simple) pragmatic inferencing work. There are interesting questions about what we predict for the processing of—or what processing might tell us about—cases where the formal criteria give conflicting results.

The final aspect to semantic complexity I will discuss has been called 'expressive complexity' by Yanovich (2012). A system A is expressively more complex than a system B if there are meanings which A can express and B cannot. Expressive complexity is different from paradigm complexity, which compares systems in which the relevant elements cover the *same* semantic space, but make more or

[11] At least in the (neo-)Gricean tradition. Some recent work argues that scalar implicatures are calculated compositionally rather than post-compositionally; see, for example, Chierchia *et al.* (2012).

[12] Jackendoff and Wittenberg (this volume, ch. 4) make a similar point when they write that "simpler grammars [. . .] put more responsibility for comprehension on pragmatics and discourse context."

fewer distinctions within that space. For example, the St'át'imcets non-future tense covers the same semantic space as the English past plus present tenses. But what if a language had no way to invoke past reference times? That would be a lack of expressive complexity when compared to either English or St'át'imcets. See Yanovich (2012) for other examples of potential ways in which languages could vary in expressive complexity.

In this section I have outlined several possible ways to measure semantic complexity, including the length of the formal description, paradigm complexity, the extent of pragmatic inferencing required, and expressive complexity. Armed with these, we now turn to the question of whether English and St'át'imcets differ in the complexity of their quantificational systems.

12.5 Is quantification in St'át'imcets less complex than in English?

Let's start by comparing the formal analyses of two particular quantifier denotations in the two languages. (19) gives a typical denotation for the English quantifier *half*, which creates a generalized quantifier. According to this denotation, *half* takes two sets, one provided by the common noun phrase and one provided by the verb phrase. It requires the cardinality of the intersection of the NP and VP sets to be half the cardinality of the NP set.

(19) $[[\ \textit{half}\ ]] = \lambda P_{<et>} . \lambda Q_{<et>} . |P \cap Q| = ½\ |P|$

As shown above, the St'át'imcets equivalents of quantifiers like *half* do not create GQs. Davis's (2010) denotation for *sáq'ulh* 'half' is given in (20).

(20) $[[\ \textit{sáq'ulh}_i\ ]]^g = \lambda P_{<et>}: |(g(i))(P)| = ½\ |P| \ . \ (g(i))(P)$

Sáq'ulh introduces a choice function, whose value is contextually given via the assignment function g. The choice function applies to the set denoted by its sister (a plural DP), and picks out a subset of that set; the entire phrase refers to the subset chosen by the choice function. The condition before the period in (20) requires the cardinality of the subset chosen to be half of that of the superset; if this fails to hold, the result is undefined.

In spite of being non-GQ-creating, the denotation for St'át'imcets *sáq'ulh* is not obviously simpler than that for English *half*. The denotation for *sáq'ulh* is not shorter, and it contains a presupposition lacking from the denotation for *half*. However, St'át'imcets *is* simpler if we adopt a metric of *type-complexity*. GQ-creating quantifiers are of type <<e,t>,<<e,t>,t>>; they take two set arguments. The St'át'imcets quantifiers in Davis's analysis are of the simpler type <<e,t>,<e,t>>, as they take one set argument, and return a smaller set. The metric of type-complexity was not discussed so far, but see for example Landman (2006) on issues of type-complexity and the drive for simpler types in certain areas of the grammar.

In terms of paradigm complexity, St'át'imcets does have a simpler quantificational system than English, because it possesses a smaller set of quantifiers than English does. As discussed by Matthewson (1998), St'át'imcets lacks DP-internal quantifiers corresponding to *most, some* or *no*. It also lacks many of the morphologically complex quantifiers that English is so good at creating, like *exactly five, fewer than five,* or *more than half.*

What about expressive complexity? We have seen that there are readings—the GQ scopal readings—that St'át'imcets quantifiers simply cannot express. Instead, only cumulative readings are possible. To be sure, there are ways of getting across the scopal readings, but they involve extensive paraphrasing. For example, a St'át'imcets speaker who wants to express wide-scope distributive universal quantification (each student read a potentially different book) can list the names of each student.[13] This need for time-consuming paraphrase suggests that St'át'imcets is less expressively complex here.

However, we mustn't forget to ask the question the other way around: does English lack readings the St'át'imcets quantificational constructions have? The answer is no and yes. Obviously, English allows cumulative readings. Scha's famous examples show that, as in (21).

(21) Exactly 600 Dutch companies use 5000 American computers.

However, at least in my judgment, cumulative readings are not easily available in English in many cases where St'át'imcets allows them. Take for example (22). In the context given, the St'át'imcets sentence is fully acceptable, but the English sentence sounds odd.

(22) *Context: Alvin and Betty were supposed to read four books each over the summer holidays. Alvin read books 1 and 2; Betty read books 3 and 4.*
[*t'qwaw's* i=sk'wemk'úk'wmi7t=a] paqwal'ikst-mín-itas [*tákem* i=púkw=a]
[*both* DET.PL=children=EXIS] read-RED-3PL.ERG [*all* DET.PL=book=EXIS]
'Both the children read all the books.'

The infelicity of the English sentence in a context which supports the cumulative reading suggests that English sentences lack readings which St'át'imcets ones have, and therefore is expressively less complex in its own way. Just like St'át'imcets does with the scopal readings, English has to paraphrase, saying something like 'The two children together read four books in total.'

What about the absence of true generics in St'át'imcets? That represents a lack of expressive complexity vis-à-vis English. As discussed above, St'át'imcets speakers have sentences which they can use in generic contexts, but the constructions they are

[13] This is similar to a strategy used in Pirahã, according to Everett (2005: 624).

forced by the grammar to use do not express fully general statements. This is similar to the situation described by Deal (2011) for Nez Perce modals. Nez Perce possesses a single non-epistemic modal *o'qa*. Deal argues that *o'qa* is a possibility modal, which can be used in necessity contexts, even though the truth conditions are not that of a necessity claim. In downward entailing contexts, *o'qa* always expresses negated possibility *(You can't go)*, and Nez Perce speakers cannot easily express negated necessity *(You don't have to go)*, having to resort to paraphrases. Deal explicitly argues that Nez Perce and English differ in effability (i.e. in expressive complexity). The same appears to be true of St'át'imcets with respect to English generics.[14]

In this section we have seen that although St'át'imcets lacks generalized quantifiers, we cannot automatically conclude that its quantificational system is less complex than that of languages which possess them. St'át'imcets quantifiers are formally less complex than English ones only if we take type complexity into account. St'át'imcets quantifier words themselves are not expressively less complex than English; the only lack of expressive complexity derives from a syntactic restriction which forces all quantifiers to co-occur with domain-restricting determiners. And St'át'imcets is paradigmatically less complex only because it happens to have a smaller inventory of quantifier words than English does. This last point has nothing to do with lacking GQs. A language could easily have only cumulative quantifiers and nevertheless possess a very large quantifier inventory.

12.6 Overall language comparison between St'át'imcets and English

In this section I compare St'át'imcets and English in a wider range of semantic subdomains. I briefly introduce the ways in which the languages differ, and discuss whether the differences reflect complexity differences. The discussion in each subsection will necessarily be brief and simplified, but the process should give us a feel for whether complexity is roughly equal between these two unrelated languages.

12.6.1 *Determiners*

St'át'imcets determiners have more paradigm complexity than English ones, in that they encode deixis, number and referentiality. The St'át'imcets determiner system is given in (23), taken from Matthewson (1998) and adapted from van Eijk (1997).

[14] Although St'át'imcets and Pirahã appear to share an absence of generics, I do not draw the same conclusions as Everett (2005) does about culture or cognition. The absence of generics in St'át'imcets does not derive from, or entail, any lack of ability to generalize, or to use intensionality or displacement. It derives from a syntactic restriction (the fact that all quantifier elements must appear within a full DP containing a determiner), and from the determiner system, which requires that all determiners encode deictic information.

(23)

	REFERENTIAL			NON-REFERENTIAL
	PRESENT	ABSENT	REMOTE	
SINGULAR	*ti...a*	*ni...a*	*ku...a*	*ku*
PLURAL	*i...a*	*nelh...a*	*kwelh...a*	

St'át'imcets does lack any equivalent of the English definite article, as all St'át'imcets determiners are indefinite (Matthewson 1998, 1999, 2009). However, this does not lead to a lack of expressive complexity in the language. Matthewson 1999 shows that in a language with only indefinite determiners, the determiners cover the same total semantic space as a language with both definite and indefinite determiners.

In terms of formal complexity, there seems to be basically a tie between the two languages, although we need the caveat that in both languages, there are a range of potential analyses to be considered. Some possible denotations for English Ds are given in (24), and two alternative denotations for the St'át'imcets singular present determiner are given in (25). At least the English definite and St'át'imcets *ti . . . a* are in the same formal complexity ballpark: both impose conditions which restrict the contexts of use, and both apply to a (possibly intensional) one-place predicate and output an individual.

(24) a. $[[\ a\ ]] = \lambda f_{<et>} . \lambda g_{<et>} . \exists x\ [f(x) = 1\ \&\ g(x) = 1]$
b. $[[\ the\ ]]^{g} = \lambda f_{<<s,e>,<s,t>>} . \lambda s : \exists! x\ f(\lambda s'.x)(s) = 1.\ \iota x\ f(\lambda s'.x)(s) = 1$ (Elbourne 2005: 51)

(25) a. $[[\ ti \ldots a_k]]^{g,c} = \lambda f_{<et>}$: (g(k))(f) is proximal to the speaker in c . (g(k))(f) (adapted from Matthewson 2001)
b. $[[ti \ldots a\ ]]^{g,c} = \lambda f_{<<s,e>,<s,t>>} . \lambda s : \exists! x\ f\ (\lambda s.x)(s_o) = 1$ where s_o is proximal to the speaker in c . $\iota x\ f(x)(s_o) = 1$ (Matthewson 2009, adapted from Elbourne 2005)

The results so far, including the results from our case study on quantifiers, are listed in (26). The language listed in each cell is the one which is more complex in the relevant area.

(26)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie

12.6.2 *Tense*

As we saw in section 12.4, St’át’imcets does not distinguish present from past tense, but has a single non-future tense morpheme, which happens to be phonologically covert. The non-future tense is shown in (27–28), and the denotation is repeated in (29). The tense picks out the contextually salient reference time, which must be non-future.[15]

(27) táyt-kan
hungry-1SG.SBJ
‘I was hungry / I am hungry / * I will be hungry.’

(28) k’ác-an’-lhkan
dry-TR-1 SG.SBJ
‘I dried it / I am drying it / * I will dry it.’

(29) [[NON-FUT$_i$]]^{g,t} is only defined if $g(i) \leq t$. If defined, [[NON-FUT$_i$]]^{g,t} = $g(i)$.

The St’át’imcets tense system is paradigmatically less complex than that of English, but equally formally complex and equally expressively complex. The results so far are given in (30).

(30)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St’át’imcets	tie
tense	tie	English	tie

12.6.3 *Modals*

The question here is whether modals in our two languages encode modality type (the difference between epistemic interpretations, deontic interpretations, and so on), and whether they encode modal force (the strength of the quantification).

English modals lexically encode modal force, but often leave modality type up to context (Kratzer 1981, among many others). This is shown in (31). The modal *must* can receive either epistemic or deontic interpretations—modality type is free—but is always interpreted as a necessity modal, with strong modal force. *May* is similarly

[15] Future time-reference is achieved by the combination of NON-FUT with a future modal *kelh,* which Matthewson (2006b) argues is an overt counterpart of English WOLL, the element which combines with tense to give *will* and *would* (Abusch 1985). I therefore take the two languages to be paradigmatically equally complex with regard to future time reference.

unrestricted with respect to modality type, but is unambiguously a possibility modal, with weak modal force.

(31) a. She *must* be in her office. EPISTEMIC / DEONTIC ONLY NECESSITY
b. She *may* be in her office. EPISTEMIC / DEONTIC ONLY POSSIBILITY

The situation in St'át'imcets is exactly reversed: The modals lexically encode modality type, but leave modal force up to context (Matthewson *et al.* 2007; Rullmann *et al.* 2008; Davis *et al.* 2009). Examples are given in (32–33), contrasting the unambiguously epistemic *k'a* with the unambiguously deontic or irrealis *ka*. Both of these modals (like all modals in the language) are felicitous in either necessity or possibility contexts; they do not form strong/weak pairs, like English modal auxiliaries do.

(32) wá7=*k'a* s-t'al l=ti=tsítcw-s=a s=Philomena
be=*EPIS* STAT-stop in=DET=house-3SG.POSS=EXIS NMLZ=Philomena
'Philomena must/might be in her house.'
ONLY EPISTEMIC NECESSITY / POSSIBILITY

(33) lán=lhkacw=*ka* áts'x-en ti=kwtámts-sw=a
already=2SG.SBJ=*DEON* see-TR DET=husband-2SG.POSS=EXIS
'You must/can/may see your husband now.'
ONLY DEONTIC NECESSITY / POSSIBILITY

In terms of paradigm complexity, the languages are roughly tied. While English does have a few more modal auxiliaries than St'át'imcets has modal clitics, this is mostly due to remnants of past-tense inflection on English modals (as in *shall* vs. *should, will* vs. *would, can* vs. *could, may* vs. *might*). As outlined above, St'át'imcets lacks a distinction between past and present tense, a fact which holds equally for its modals. But in terms of modality type and modal force, the languages are equal in complexity, as indicated in the simplified tables in (34–35). Each language chooses one distinction to encode and one distinction to leave up to context. These tables also show that the languages have equivalent expressive complexity, in that the same total semantic space is covered.

(34)

English	NECESSITY	POSSIBILITY
DEONTIC	*must*	*may*
EPISTEMIC	*must*	*may*

(35)

St'át'imcets	NECESSITY	POSSIBILITY
DEONTIC	*ka*	*ka*
EPISTEMIC	*k'a*	*k'a*

In order to measure formal complexity, we as usual need analyses of both systems. Versions of these are given in (36–39). The first argument of the English modals is a conversational background function from worlds to sets of worlds; the conversational background is given by the context and determines the modality type.[16]

(36) $[[\ \textit{might}\ ]]^{c,w} = \lambda g_{<s,st>}.\ \lambda p\ .\ \exists w' \in g(w)\ [p(w') = 1]$

(37) $[[\ \textit{must}\ ]]^{c,w} = \lambda g_{<s,st>}.\ \lambda p\ .\ \forall w' \in g(w)\ [p(w') = 1]$

In St'át'imcets, the conversational background is lexically restricted to being either epistemic or deontic. All modals are necessity modals (i.e. they introduce universal quantification over worlds), but variable modal force is achieved by a choice function which narrows down the set of worlds being quantified over. The lexical entry for *k'a* is given in (38). *Ka* is exactly the same except the word 'deontic' replaces 'epistemic'. See Rullmann *et al.* (2008) for details.

(38) $[[\ \textit{k'a}\ ]]^{c,w}$ is only defined if c provides an epistemic conversational background g and a modal choice function f such that $f(g(w)) \subseteq g(w)$.
If defined, $[[\ \textit{k'a}\ ]]^{c,w} = \lambda f_{<st,st>}.\ \lambda g_{<s,st>}.\ \lambda p\ .\ \forall w' \in f(g(w))\ [p(w') = 1]$
(adapted from Rullmann *et al.* 2008)

The St'át'imcets denotations are formally more complex than the English ones. The St'át'imcets modals take more arguments (as a modal choice function is required), and the St'át'imcets modals have presuppositions not present in English. An interesting thing to think about, however, is whether the complexity of the St'át'imcets denotations has partly to do with the way predicate logic is constructed. Predicate logic gives us single symbols for both universal and existential quantification—thus making the English denotations simple—but does not provide single symbols for either epistemic or deontic modality. This adds to the point made in section 12.4 regarding the difficulty inherent in measuring formal complexity by a length criterion.

The results so far are summarized in (39).

(39)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie
tense	tie	English	tie
modals	St'át'imcets	tie	tie

[16] In the Kratzerian system (Kratzer 1981), there are actually two conversational backgrounds, a modal base and an ordering source. I am ignoring this here for the sake of simplicity.

12.6.4 *Evidentials*

Evidentials encode the type of evidence the speaker has for the claim expressed. St'át'imcets possesses a set of these, as shown in (40); see Matthewson *et al.* (2007), Matthewson (2012) for analysis.[17]

(40) *k'a* inferential (any kind of indirect evidence)
an' perceived evidence (sensory witness of results)
lákw7a sensory non-visual
ku7 reportative

Although English epistemic modals may all convey the evidential notion of 'indirect evidence' (von Fintel and Gillies 2010; Matthewson, in press), they do not encode sub-distinctions between different types of evidence. St'át'imcets is therefore both formally and paradigmatically more complex in this area. Expressively, there is no difference; English modals are compatible with all types of indirect evidence, and therefore the same total semantic space is covered by English epistemic modals and St'át'imcets evidentials.

(41)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie
tense	tie	English	tie
modals	St'át'imcets	tie	tie
evidentials	St'át'imcets	St'át'imcets	tie

12.6.5 *Presuppositions*

A presupposition is a proposition which is 'taken for granted' at the time the sentence containing it is uttered (Stalnaker 1974). Presuppositions thus impose contextual felicity constraints. If a presupposition is not common knowledge at the time of utterance, the hearer may express surprise by means of responses like 'Hey, wait a minute! I didn't know that…' (von Fintel 2004). In English, certain items lexically encode presuppositions; examples include the definite article, and words like *again, too,* and *stop.*

[17] St'át'imcets evidentials encode not only evidence source, but also epistemic modality (Matthewson *et al.* 2007). Hence, the modal *k'a* discussed in the previous section is included in (40).

St'át'imcets has been argued to lack presuppositions which impose contextual felicity constraints in this way (Matthewson 2006a, 2009). Evidence for this claim includes the failure in St'át'imcets of the 'Hey, wait a minute!' test; there are no elements in St'át'imcets which induce surprise responses when some aspect of their meaning fails to be taken for granted by the hearer. St'át'imcets does have versions of words like *again, too,* and *stop,* and the relevant parts of their meanings share with English presuppositions the property of being not-at-issue (being outside the main asserted content of the utterances; Potts 2005; Tonhauser *et al.* 2013, among others). In other words, St'át'imcets possesses only a subset of the kinds of not-at-issue meaning available in natural language, namely, the non-presuppositional kind. These non-presuppositional not-at-issue meanings correspond to Potts's (2005) definition of conventional implicatures.

Which is more formally complex: a presupposition, or some other kind of not-at-issue content? The answer is not clear, especially since the field has not reached agreement on how non-presuppositional, not-at-issue content should be modeled. Potts places conventional implicatures in a separate dimension of meaning, which would seem to add formal complexity, but the jury is still out on many issues surrounding multi-dimensionality.

The question of paradigm complexity does not really arise here, since the languages do not differ in the number of elements, but only in whether the elements impose contextual felicity constraints. What about expressive complexity? This is actually an interesting case. On the one hand, St'át'imcets seems to lack expressive complexity, because it lacks an entire class of meanings. On the other hand, suppose we take any particular translation pair (such as English *again* and St'át'imcets *múta7* 'again'). In all cases where it *is* common knowledge that an event has happened before, both presuppositional *again* and non-presuppositional *múta7* are felicitous,

(42)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie
tense	tie	English	tie
modals	St'át'imcets	tie	tie
evidentials	St'át'imcets	St'át'imcets	tie
presuppositions	?	English	tie

and no difference between them should be detectable. In cases where the relevant information is *not* already known to the hearer, St'át'imcets *múta7* will be straightforwardly felicitous, and English *again* should either require presupposition accommodation, or will require paraphrases which lose the not-at-issue status of the information (e.g. 'This has happened before and . . . '). In the last case, it is English which lacks expressive power compared to St'át'imcets.

In the absence of any clear winner in expressive complexity here, I'm going to call it a tie.

12.6.6 *Argument structure and control*

In English, the valency of predicates is not overtly encoded; both (43a) and (43b) contain the same verb.

(43) a. Roland wrote. INTRANSITIVE
b. Roland wrote a story. TRANSITIVE

Not so in St'át'imcets, which obligatorily marks (in)transitivity on every predicate, as shown in (44).

(44) a. mets-*cál* kw=s=Roland
write-*INTR* DET=NMLZ=Roland
'Roland wrote.' INTRANSITIVE

b. mets-*en*-ás kw=s=Roland ti=sqwélqwel'=a
write-*TR*-3ERG DET=NMLZ=Roland DET=story=EXIS
'Roland wrote a story.' TRANSITIVE

The St'át'imcets (in)transitivizers also encode (roughly) whether the subject argument is an agent in full control of the event (van Eijk 1997; Davis 2012). This is shown in (45), with the transitivizers glossed more fully to indicate their control status.

(45) a. zík-*in'*-as ta=sráp=a i=sám7=a
fall-*CONTROL.TR*-ERG DET=tree=EXIS DET.PL=white.person=EXIS
'The white people cut the tree down (deliberately).' (Davis 2012: ch. 39)

b. zikt-*s*-ás ta=sráp=a ts7a ku=xwélmen
fall-*NON.CONTROL.TR*-ERG DET=tree=EXIS DEM DET=saw
'This saw cut the tree down.' (Davis 2012: ch. 39)

c. zík-*am* i=sám7=a
fall-*CONTROL.INTR* DET.PL=white.person=EXIS
'The white people did some tree-cutting.'

St'át'imcets is both formally and paradigmatically more complex in this area than English. Expressively, there is no difference; both languages have transitive and intransitive verbs, and can use them in cases where the subject is in control, and when s/he is not. The final summary of results is in (46).

(46)

	FORMAL COMPLEXITY	PARADIGM COMPLEXITY	EXPRESSIVE COMPLEXITY
quantifiers	tie	English	English
determiners	tie	St'át'imcets	tie
tense	tie	English	tie
modals	St'át'imcets	tie	tie
evidentials	St'át'imcets	St'át'imcets	tie
presuppositions	?	English	tie
argument structure	St'át'imcets	St'át'imcets	tie
control	St'át'imcets	St'át'imcets	tie

Of course, I have not presented a complete language comparison, since I have looked only at a small sub-part of each language, and have set pragmatic inferencing complexity aside due to the embryonic state of our knowledge in this area. From what we have looked at here, St'át'imcets is winning the complexity race on points.

It is perhaps more interesting, however, to think about the way complexity is distributed across the different ways of measuring. In the area of formal complexity, St'át'imcets is the clear winner, but the only clear expressibility difference involves English being more complex. With respect to paradigms, each language seems to concentrate its complexity in a different subset of areas. While English has a more complex quantifier system, St'át'imcets has a more complex determiner paradigm, and so on. From a functional standpoint, this distribution of complexity is probably not a coincidence. That is, we might expect that languages gravitate towards a roughly equal overall number of paradigmatically complex areas. This leads to the impression that languages 'care about' different semantic distinctions. For example, some languages 'care about' evidence source and have rich evidential paradigms; these are usually disjoint from the languages which 'care about' modal quantificational force and have paradigmatically rich modal systems which encode strength distinctions.

12.7 Conclusions

The first lesson from our case study on English and St'át'imcets was that determining whether languages differ in their semantics is a task which requires targeted, hypothesis-driven fieldwork. We then saw that even once we have established how the languages in question vary, we should still not leap to conclusions about complexity. The fact that a language lacks some set of meanings which another language possesses (such as generalized quantifiers) does not automatically mean the former language is more simple. I have argued that in semantics, complexity must take into account not only formal complexity and paradigm complexity, but also pragmatic inferences, and the range of meanings a language can achieve (expressive complexity). We saw that while St'át'imcets noun phrases do not participate in scopal ambiguities, they freely allow cumulative readings in situations where English noun phrases do not, and are thus in one sense expressively more complex than their English counterparts. However, St'át'imcets does lack real generic readings, an absence which represents an expressive complexity difference with English.

The fact that languages vary in expressive complexity is a non-trivial result. It means that Fromkin *et al.* were wrong to say (2010: 34) that "All languages are equally complex and equally capable of expressing any idea," and it means that Katz's (1976: 37) Strong Effability Hypothesis, that "every proposition is the sense of some sentence in each natural language," is false. Expressive complexity is intuitively the most radical way in which languages can differ, and at the same time it is the type of complexity the human language faculty is least likely to be able to constrain. That is, Universal Grammar does not plausibly contain a list of meanings which all languages must be able to express. The weak expressive equivalence which does exist probably derives from general cognition and our status as social and cultural beings. Universal Grammar places formal and substantive constraints on meaning, but not on semantic complexity per se. So some languages could be, overall, semantically more complex than others.

13

Computational complexity in the brain

CRISTIANO CHESI AND ANDREA MORO

13.1 Introduction

How complex is (a) human language?[1] Every chapter in this volume is devoted, in one way or another, to addressing this question. The great majority of the chapters focus on the question of whether languages can be *compared* in terms of their relative degree of complexity. The present chapter, however (along with Trotzke and Zwart, chapter 7) takes a somewhat different approach to the question of grammatical complexity. We ask: 'How might one characterize the degree of complexity of natural language *in general*?' Fortunately there are formal means of measuring the complexity of grammatical systems that go back to the 1950s. We refer, of course, to the Chomsky hierarchy (Chomsky 1956, 1963), which classes grammars, and the automata that recognize them, in formal terms. Our chapter goes beyond that of Trotzke and Zwart, however, in attempting to identify the brain activity that correlates with the formally-arrived at levels of grammar in the hierarchy. We will see that the relationship between the hierarchy and both measurable processing difficulty and detectable brain activity is not a simple one.

The chapter is organized as follows. Section 13.2 reviews briefly the Chomsky hierarchy. Section 13.3 investigates the processing (primarily parsing) correlates of the hierarchy and section 13.4 the neurophysiological correlates. Section 13.5 further probes the modeling by automata of neurolinguistics processes. Section 13.6 is a brief conclusion.

[1] We wish to thank Theresa Biberauer, Barbara Citko, Ray Jackendoff, Fritz Newmeyer, Laurel Preston, Ian Roberts, and the participants at the workshop on Formal Linguistics and the Measurement of Grammatical Complexity for helpful discussion. Special thanks go to Fritz Newmeyer and Laurel Preston, for careful review and insightful comments on previous drafts of this chapter.

13.2 The Chomsky hierarchy

In this section we provide background on the definition of formal grammars, the ranking of languages generated by such grammars as given by the Chomsky hierarchy, and the position of natural languages with respect to the hierarchy. We then relate the generative power of the grammar types on the hierarchy to computational procedures that can be used to calculate complexity costs, through the use of automata.

A formal grammar consists of a finite set of production rules (left-hand side → right-hand side), where each side consists of a sequence of the following symbols: a finite set of nonterminal symbols (indicating that some production rule can yet be applied), a finite set of terminal symbols (indicating that no production rule can be applied), and a start symbol (a distinguished nonterminal symbol). A formal grammar defines (or generates) a formal language, which is a (usually infinite) set of finite-length sequences of symbols (i.e. strings) that may be constructed by applying production rules to other sequences of symbols which initially contain just the start symbol.

The Chomsky hierarchy consists of the following levels:[2]

(1) a. Type 0: unrestricted (or Turing-equivalent) grammars. All rules are of the form $\alpha \rightarrow \beta$, with the only restriction being that α cannot be null.
b. Type 1: context-sensitive grammars (CSGs). All rules are of the form $\alpha A\beta \rightarrow \alpha\gamma\beta$. A must be nonterminal and α, β, and γ can be strings of terminals and nonterminals. The strings α and β may be empty, but γ must be nonempty.
c. Type 2: context-free grammars (CFGs). All rules are of the form $A \rightarrow \gamma$. Only one non-terminal symbol is allowed to the left side of the rewriting rules.
d. Type 3: regular grammars (RGs): All rules are of the form $A \rightarrow xB$ or $A \rightarrow Bx$, where x is a terminal symbol (i.e. a lexical item). Note that any non-terminal symbol must appear systematically either on the right edge (right RGs) or to the left edge (left RGs).

Figure 13.1 depicts the set inclusions of languages according to the hierarchy. Grammars with fewer restrictions can generate all the languages that are generable by grammars with more restrictions.

Note the positioning of 'natural languages'. It has been argued that they require a grammar with a generative power that exceeds that of RGs (Chomsky 1957) and of CFGs (Shieber 1987). The arguments supporting this claim are based on two kinds of recursive structures that are attested in natural languages. Consider the following examples:

[2] The remainder of this section greatly oversimplifies the material treated. The interested reader should consult any introduction to mathematical linguistics for more refined detail.

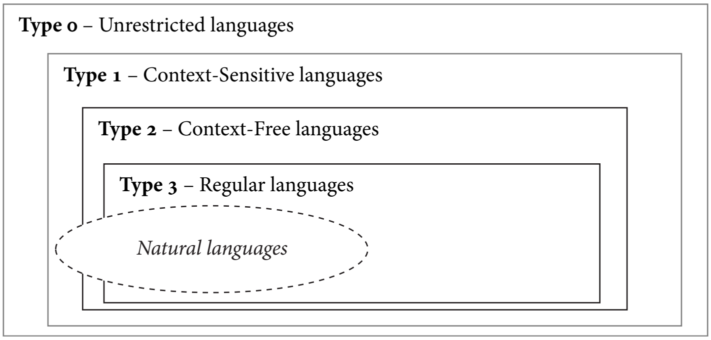


FIGURE 13.1 Set inclusions in the Chomsky hierarchy.

(2) a. The rat$_1$ [(that) the cat$_2$ [(that) the dog$_3$ chased$_3$] ate$_2$] died$_1$

b. Anyone$_1$ [who feels that if$_2$ [so-many$_3$ [more$_4$ [students$_5$ [whom we$_6$ haven't$_6$ actually admitted] are$_5$ sitting in on the course] than$_4$ ones we have] that$_3$ the room had to be changed], then$_2$ probably auditors will have to be excluded], is$_1$ likely to agree that the curriculum needs revision.

(Chomsky and Miller 1963: 286)

c. Alberto, Bianca ... e Xenia sono rispettivamente promosso, bocciata ... e promossa. (Italian)

A. $_{male}$, B. $_{fem}$ ... and X. $_{fem}$ are respectively promoted $_{male}$, rejected $_{fem}$ ... and promoted $_{fem}$

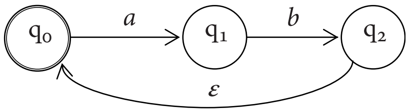
(2a–b) represent cases of nested dependencies of the 'mirror recursion' type XX^R (e.g. *abcddcba*), while (2c) illustrates a case of 'cross-serial dependency' of the XX kind (e.g. *abcdabcd)*. Partee *et al.* (1990) show how sentences containing nested dependencies cannot be generated/recognized by RGs and sentences containing cross-serial dependencies cannot be generated/recognized by CFGs. Given such data, it appears to be the case that natural language grammars have at least the generative power of context-sensitive grammars.[3]

Our goal now is to relate generative power as ranked by the hierarchy to computational procedures for which complexity costs can be calculated. One way to evaluate complexity, relying on the Chomsky hierarchy, is to use automata that are equivalent in terms of generative power to the grammars in the hierarchy (for discussion, see Hopcroft *et al.* 2001). Finite state automata (FSAs), for instance, can express any regular language by using states and transitions among states. The diagram below shows how simple recursive rules characteristic of RG, such as (3a),

[3] Or perhaps mildly context-sensitive grammars (MCSG) (see Weir 1988 for discussion). MCSG express a class of grammars not included in CFGs, but included in CSGs (the tree-adjoining grammars of Joshi 1985 are an example).

can be subsumed by finite automata like the one in (3b). Another way of putting it is to say that (3a) and (3b) capture the same set of sentences:

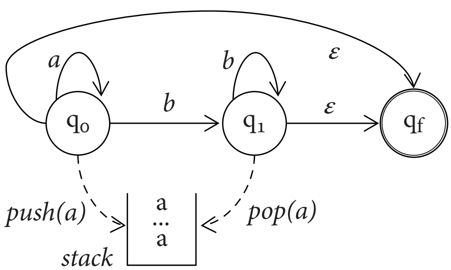
(3) a. S $\rightarrow$ *a b* S
S $\rightarrow$ ε
b.



The automaton starts in q_o (the initial state, by convention, is labeled q_o) and ends in q_o (in this special case, the initial state is also the final state; in the diagrams, by convention, we indicate final states using double-circles). The arrows indicate possible transitions from one state to the other and they can be traversed only when the (terminal) symbol (e.g. a lexical item) that labels them is recognized or generated (ε indicates the null symbol; the transition labeled by such a symbol can be freely traversed without consuming input tokens). Only if a final state is reached, and every token in the input is consumed, can the computation terminate successfully.

CFGs are more powerful than RGs and hence require more powerful automata. The latter are called push-down automata (PDAs). The example below depicts a sample CFG and a corresponding automaton:

(4) a. S $\rightarrow$ *a* S *b*
S $\rightarrow$ ε
b.



Both CF rewriting rules like those in (4a) and their corresponding automata (4b) can capture a counting dependency (that is formally equivalent to a mirror recursion) of the $a^n b^n$ kind (i.e. *n* sequences of *a* followed by the very same number of *b*). This automaton will start its computation from the initial state q_o, then it will 'push' an *a* in the stack any time it reads one from the input. Reading a *b*, the automaton will move to q_1 and it will start 'popping' an *a* as long as *b*s are in the input. Once *b*s are exhausted, if the stack is empty, the automata will reach q_f, the final state, signaling the end of a successful computation. The stack in PDA hence has a 'last in first out' (LIFO) memory structure. That is, the last item pushed in the stack will be the first one to be popped out. The computation starts at q_o, the initial state, and ends at q_f, the final one.

Now then, where does the issue of 'complexity' enter the picture? Given the above, we can conceptualize complexity in two dimensions: time and space. The time complexity is a function of the number of states traversed. As can be seen in comparing (3b) and (4b), FSAs and PDAs with the same number of states are equivalent in terms of state traversals: given a string of length 4 (*abab* for the FSA, *aabb* for the PDA), both automata need 4 state traversals (the initial states and the states reached using the ε transition are not counted). However, if we take a 'three dimensional' look at the diagrams, we can see that processing *aabb* using the PDAs is more complex than processing *abab* using the FSAs in terms of space complexity. The space complexity is a function of the additional memory needed to keep track of the dependency. A PDA is thus more complex than a FSA in a very precise way, namely it needs a LIFO memory buffer. So to process a string of length 4 (i.e. *aabb*), in addition to the 4 state traversals, the PDA needs to store 2 items in the memory (the two instances of *a*), using 4 operations (pushing *a* twice and popping *a* twice). For a string of length 6 (i.e. *aaabbb*) the very same PDA needs 6 state traversals, 3 items in the memory buffer, and 6 operations on the memory buffer.

For our purposes, the important thing to keep in mind while reading the following section is that such computational models give us a baseline for comparing precisely the complexity of two different performance tasks (such as processing different levels of hierarchical embedding) in terms of time (the number of computational states to be traversed) and space (the number of items to be stored and retrieved).

13.3 Computational models of processing complexity

This section takes as its point of departure the idea that the Chomsky hierarchy translates directly into a measure of processing complexity. However, we will see that this is not correct. Many factors affect processing that are independent of the position of a grammar on the hierarchy. We will discuss three examples where complexity as measured by the hierarchy does not directly correspond to complexity as measured by processing tasks. We will then discuss additional factors that may account for this lack of direct correspondence.

Various hypotheses have been discussed in the literature that propose a more or less straightforward correlation between the automata (and their relative complexity metrics) discussed in the previous section and the degree of difficulty in human sentence processing. For example, it has been proposed that patterns that are not captured by RGs are generally hard to process (Pullum and Gazdar 1982: 498). One might assume then that cross-serial dependencies, such as (2c), which are licensed by the more complex CFGs and CSGs, should be even more difficult to parse and generate than nested dependencies such as (2a–b). This is roughly, but not entirely, correct. Bach *et al.* (1986) compared, in terms of acceptability judgments and accuracy in paraphrase understanding, the linguistic performance on cross-serial and nesting dependencies.

They showed that cross-serial patterns (in Dutch) with two levels of dependency are significantly more easily processed than center embeddings with the very same level of dependencies (in German).[4] Comparable results have been obtained with three levels of dependency using artificial grammars, both with humans and with artificial neural network simulations (Christiansen and Chater 1999). In other words, what seems crucial in determining processing complexity is the kind of dependency, with nesting dependencies in general harder than cross-serial dependencies. The position in the Chomsky hierarchy is not the determining factor.

Another way to relate the Chomsky hierarchy to processing complexity involves looking at specific performance tasks. For example, it has been suggested that performance difficulty in parsing is associated with the limitations of the memory stack (Yngve 1960). Given such an idea, the more incomplete (i.e. unexpanded) the phrase structure rules are that the parser needs to store within the stack (to retrieve a complete phrase structure), the more complex the sentence will be. In this vein, Chomsky and Miller (1963) suggested that a stack-based parser might simply get confused while processing self-embedding. The problem is that storing more than once the same context-free rules in the stack might lead the parser to confuse two instances of the very same rule. This problem holds only for CF (or generatively more powerful) rules and crucially not for RG rules, since only in the first case is a new rule stored in the stack while the previous rule is not yet completed. In RGs, top-down expansion of any rule does not need any stack-based storage.

Another problem is that sentences with the same level of embedding (and otherwise equivalent in terms of the Chomsky hierarchy) are perceived as more or less difficult to process depending on the position of the very same lexical items. It seems to be the case that the complexity of certain structures is not related simply to the phrase structure to be expanded or to the necessity of using a memory buffer, but to the kind of non-local dependency to be computed. For instance, consider the sentences in (5) below. The restrictive relative clause whose head (*reporter*) is interpreted in the object position within the relative clause, hence an 'object-headed relative clause' (ORC; 5a), takes more time to process than a subject-headed relative clause (SRC; 5b) (King and Just 1991):[5]

(5) a. The reporter$_i$ [who [the senator] attacked _$_i$] admitted the error. (ORC)
    b. The reporter$_i$ [who _$_i$ attacked [the senator]] admitted the error. (SRC)

[4] German and Dutch are compared since the equivalent sentence requires a nested dependency [arg$_1$ [arg$_2$ verb$_2$] verb$_1$] in German and a serial dependency in Dutch [arg$_1$ arg $_2$ verb$_1$ verb$_2$].

[5] This asymmetry has been systematically documented using self-paced reading experiments (King and Just 1991), probe-task paradigms (Wanner and Maratsos 1978), brain activity analysis (Just *et al.* 1996), and eye-tracking techniques (Traxler *et al.* 2002).

What seems at issue in (5) is the special relation between the filler and the gap (Fodor 1978). In particular, the kind of element that intervenes (*the senator* in 5a) between the filler (*the reporter)* and the gap (the argument position within the relative clause) appears to play a crucial role.

Pursuing the question of the processing costs of the two types of relative clauses, it has been found that pronouns and proper names produce a milder complexity increase in sentence processing than full DPs in ORCs. Syntactic prediction locality theory (SPLT, Gibson 1998) attempts to explain why this is the case. Gibson's idea recasts complexity metrics in terms of *integration* and *memory-load cost*. The first component (integration cost) is associated with new discourse referents to be incorporated in the structure (a full DP introduces a new referent, a pronoun does not). The second component (memory-load) is associated with keeping track of obligatory syntactic requirements, that is, the minimal set of items that could conclude the processed sentence in a well-formed way. For instance, given independently-needed assumptions about the grammar-processor relation, after processing a DP like *the reporter*, the processor expects at least a VP (e.g. *left*) to complete the sentence. Waiting for a VP will have a cost that needs to be added to the cost of other expectations plus the cost of integrating new discourse referents, if any. In (5a), *the senator* introduces a new discourse referent and therefore incurs an integration cost. (5a) incurs a memory-load cost because the processor must wait for the VP until after the relative clause is complete.

PDAs appear to be computationally adequate for implementing both integration cost and memory-load cost. As far as the former is concerned, we might assume that the recognition of any new referent (a proper name or a full DP) triggers the insertion of a co-indexed variable in the memory buffer, while a pronoun picks up a referent from it (cf. Schlenker 2005; Bianchi 2009). As for memory-load cost, we might assume that any new sequential DP moves the automaton to a deeper state more removed from the final state. If this were correct, then integration cost and the memory-load would translate directly into space complexity and time complexity respectively. But things, as always, are more complicated than that. Unfortunately, Gibson's hypothesis makes incorrect empirical predictions where intervening proper nouns are involved. Consider (6a–c):

(6) a. *The pictures* that *the boy* took yesterday
    b. *The pictures* that *John* took yesterday
    c. *The pictures* that *I* took yesterday

(6a) is measurably harder to process than (6b), which is harder than (6c) (Gordon *et al.* 2001, 2004). To address this problem of incorrect prediction, Warren and Gibson (2005) have refined Gibson's integration cost by defining a hierarchy of referents, with full DPs 'heavier' than proper names, and proper names 'heavier' than pronouns. Doing so explains the contrast schematized by the paradigm in

(6), but not the fact that when the relative clause head is also a proper name (*It is **Bob** that **John** saw yesterday*), the processing difficulty is comparable to that of (6a) (Belletti and Rizzi 2013).

These contrasts suggest that factors other than referentiality are involved in the correct complexity metric for processing. Among these are animacy features (Mak *et al.* 2002) and agreement features (Belletti *et al.* 2012). As far as features are concerned, locality constraints (Rizzi 1990; Friedmann *et al.* 2009) seem to play a fundamental role in accounting for complexity increase in terms of feature intervention. We can rephrase part of Friedmann *et al.*'s (2009) hypothesis as follows:

(7) Feature intervention: When *X* and *Y* enter a non-local relation and *Z* intervenes within such relation, the complexity of the dependency is proportional to the features shared between *X* and *Z*.

We could adapt (7) to a model of processing cost by modifying Gibson's idea of integration cost in the following way:

(8) Feature-based integration cost (FIC, revision of Gibson 1998's model of integration cost):
Any operation in the memory buffer has a cost that is proportional to the number of features the item that must be stored or retrieved shares with other items already present in the memory buffer.

We assume that in ORC like (6a), the head of the relative clause, *the pictures*, which is the filler to be stored in memory, is expressed in terms of features as [+D +N] (i.e. determined, nominal entity), as is *the boy*. Therefore two features are shared and two operations (storage and retrieval from memory) are needed for each of them. We combine the two factors in one naïve complexity function C_{FIC} as follows:

(9) $C_{FIC} = O^F$ (naïve Feature-based Integration Cost function)

O is the number of operations accessing memory (two in this case, namely storage and retrieval) and F is the number of shared features (also two in this case).

Note that that this rather simple complexity function (see Chesi 2014 for an elaboration of this) automatically explains the contrasts in (6a–c). Assuming that what distinguishes a common noun from a proper noun is a sub-specification of one single nominal feature (i.e. N vs. N_{proper}, as in Belletti and Rizzi 2013), we could consider this overlapping as half of the cost of a full feature overlapping (i.e. 0.5). Hence in (6a), the FIC for *the boy* will be $C_{FIC} = 4$ (i.e. 2^2), while in (6b), *John* will have a $C_{FIC} = 2.8$ (i.e. an approximation of $2^{1.5}$, since +D is fully shared but N and N_{proper} are counted as half sharing). Finally, in (6c) *I* will cost $C_{FIC} = 2$ (since only +D is shared, assuming that a pronoun lacks the N feature (Friedmann *et al.* 2009). Hence, (8) not only explains the processing asymmetry revealed in (6a–c), but also the fact that when the head of the relative clause and the filler are both proper nouns, (*it is* ***Bob*** *that* ***John*** *saw yesterday*) we get a C_{FIC} for *John* of 4, exactly as in (6a).

To summarize, we have seen that patterns that are captured by more powerful grammars in the Chomsky hierarchy are not always more complex to process than patterns that require less powerful grammars. Independently of the hierarchy, the constructions discussed in this section require the specification of at least two components for a correct complexity function characterization: a hierarchical component and a component that encodes features involved in non-local dependencies of the filler-gap type. We turn now to the question of complexity as revealed by operations in the brain.

13.4 Syntax in the brain: autonomy, hierarchy and locality

Research on the biological foundations of language has seen a dramatic development since the turn of the century (see Cappa 2012; Friederici 2011; Kandel *et al.* 2012 for reviews). In particular, the possibility of exploiting neuroimaging techniques has offered interesting opportunities for deepening our understanding of the relationship between syntax and the brain. In this section, we discuss whether the main components of complexity discussed in section 13.3, hierarchy (time complexity, i.e. level of structural embedding) and locality (space complexity, i.e. memory) are distinguishable, not only at the computational level, but also at the neurological level. The following two subsections develop the grammar–brain relationship. Section 13.4.1 reviews evidence that syntactic computation activates a dedicated network and takes on the question of where in the brain syntax is activated. Section 13.4.2 discusses the complexity signatures of hierarchical syntactic processing and non-local dependency formation.

13.4.1 *Where to look for complexity: syntactic networks in the brain*

The classical assumption is that the language centers in the brain are Broca's and Wernicke's areas. However, imaging techniques such as PET and fMRI have shown that the story is not quite that simple.[6] On the one hand, additional areas are involved in language processing (some of them deeply brain-internally), on the other hand, both Broca's and Wernicke's areas are more finely subdivided according to their specialization revealed in distinct functions.

To begin, isolating syntax from other linguistic components by pointing to a difference in neural activity is not an easy task, since processing does not treat syntax

[6] A PET scan is an imaging technique that exploits (fludeoxy)glucose uptake to reveal neural metabolic activity by tracing regional concentration of these molecules. PET has a good spatial resolution, but a poor temporal resolution. An fMRI is an imaging technique that allows us to visualize neural activity by detecting associated changes in blood flow by using blood-oxygen-level-dependent contrast detectors. fMRI has an excellent spatial resolution and a decent temporal resolution.

as an isolated entity. Moro *et al.* (2001) tackled this problem by inventing a lexicon consisting of fake nouns, verbs, adjectives, etc., whose entries were phonotactically compatible with Italian. Real Italian functional morphemes (determiners, inflections, etc.) were combined with these fake content words, thereby creating pseudo-Italian sentences. The experimental stimuli consisted in introducing selective errors at different levels in 'quasi-Italian' sentences of this kind (for ease of exposition we replace them with 'quasi-English' sentences):

(10) a. The gulk ganfles the brals (grammatical sentence)
b. The gulk ganfrzhrld the brals (phonological error)
c. The gulks ganfles the brals (morphosyntactic error)
d. Gulk the ganfles brals the (syntactic error)

The subjects were prompted with baseline stimuli (well-formed pseudo-sentences like 10a) and their brain activity was compared during the processing of sentence with errors of various types (10b–d). The subjects were asked to read silently the visually presented sentences and, for the three error types under analysis, to provide an acceptability judgment. Brain activity was recorded using a PET (Positron Emission Tomography) scan.

Interestingly, the results show that syntactic anomalies activate specific brain regions, in particular Broca's areas BA44 and BA45 and the Left Nucleus Caudatus (LNC, part of the basal nuclei in the inner brain). Figure 13.2 presents a schematic of brain areas and their possible functions. Further evidence in support of Figure 13.2 will be discussed below.

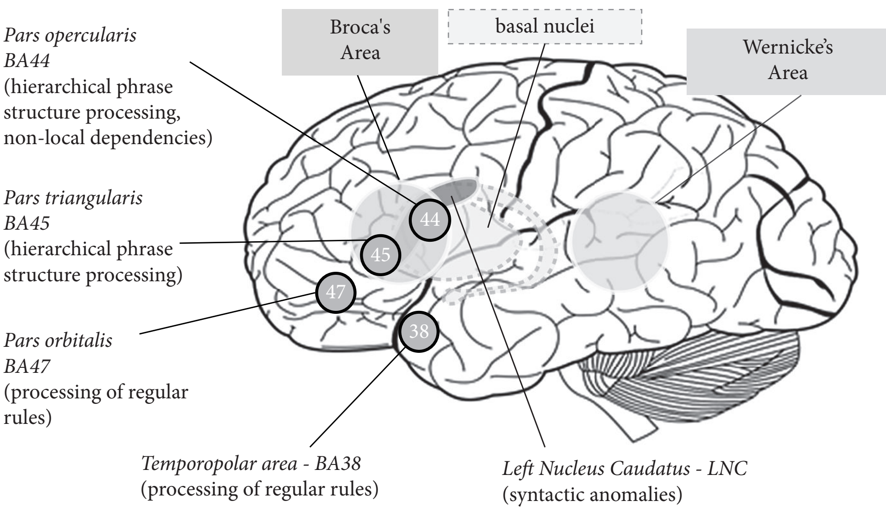


FIGURE 13.2 A very schematic summary of brain locations (and their possible functional roles).

These results are compatible with other studies that have attempted to isolate syntactic processing networks. For instance Embick *et al.* (2000), also using a violation task, found a selective activation of Broca's area in grammatical vs. orthographic errors. This latter experiment, however, did not exclude the lexical semantic contribution and thus provided a weaker result, in that it was possible that syntactic errors also activated semantic networks.

Before concluding that BA44/45 and LNC are the (only) places to look for syntactic complexity, we should take into consideration a few more studies that greatly refine our understanding of BA44/45 and LNC. For example, Musso *et al.* (2003) provided evidence that the distinction between possible (i.e. Universal Grammar-compliant) and impossible (i.e. Universal Grammar-unrealistic) rules is not just an artificial taxonomy resulting from cultural or conventional constructs, but is instead the reflex of a genuine neuropsychological structure. In order to test this hypothesis, an fMRI experiment was designed to compare syntactic rules sensitive to hierarchical constituent structure to impossible rules targeting a fixed position in the sentence. German speakers from strictly monolingual communities in the pre-unification German Democratic Republic participated in the experiment. Subjects were asked to learn 'artificial grammars' of a pseudo-Italian language, that is, those composed of sentences containing Italian lexical items, but generated either by real Italian rules or by fake (impossible) rules.[7] In the experiment in which real Italian rules were used, the experimental subject would be expected to invert the grammatical subject and the object positions to create a passive (11a). In another, he or she would be expected to place the complementizer *che* between the matrix and subordinate clause. In another experiment, impossible rules were used. For instance, to correctly complete the task the participant would be expected to place the negation marker after the third word (12a) or to force agreement between the first and the last word in the sentence (12b). Note that all rules apply to the baseline sentence *Paolo mangia la pera* ('Paolo eats the pear').

(11) Possible rules

a. **La pera** è mangiata **da Paolo** (passivization)
the pear is eaten by Paolo

b. Pia dice **che** Paolo mangia la pera (complementation)
Pia says that Paolo eats the pear

(12) Impossible rules

a. Paolo mangia la **no** pera (negation)
Paolo eats the no pear

b. Un**a** bambino mangia una per**a** (agreement)
$\text{a}_{\text{fem.sing}}$ child eats a $\text{pear}_{\text{fem.sing}}$

[7] See Moro (2008) for a discussion of the characterization of 'possible' and 'impossible' rules.

Interestingly, it was only the processing of possible rules that correlated with stronger activation of part of Broca's area (pars triangularis, BA45), although all subjects rapidly acquired the same capability to manipulate both possible (11a–b) and impossible (12a–b) rules. Hence we have robust evidence in favor of the neurobiological reality that only hierarchical (syntactic) rules activate a specific brain region and that human language is not purely a matter of cultural or arbitrary convention.[8]

Other studies bear out the idea that BA44/45 is involved in hierarchical and/or non-local processing of linguistic stimuli, for example, those involved in counting dependencies of the $a^n b^n$ kind, which require, as we have seen, the generative power at least of CFGs or PDAs (see Fitch and Friederici 2012). However, patterns that do not activate BA44/45 include strings that conform to the pattern $(ab)^n$, namely sequences of any length of couples of items like the syllables *pa* and *do*: *pa-do-pa-do-pa-do*. . .). This pattern selectively activates brain regions BA38 (the temporopolar area) and BA47 (part of the inferior frontal gyrus, also known as pars orbitalis).

The hypothesis that only the activation of BA44/45 is selective to hierarchy and/or non-local dependencies is also consistent with the fact that a selective activation of the posterior part of Broca's area has been revealed during tasks comparing syntactic vs. purely lexical (semantic) conditions. On this point, Dapretto and Bookheimer (1999) showed, in a fMRI study, that BA44 is more active when syntactic relations are involved (13a–b), as opposed to purely lexical ones (14a–b):

(13) a. The policeman arrested the thief.
a'. The thief was arrested by the policeman.

b. The teacher was outsmarted by the student.
b'. The teacher outsmarted the student.

(14) a. The lawyer questioned the witness.
a'. The attorney questioned the witness.

b. The man was attacked by the Doberman.
b'. The man was attacked by the pit bull.

Earlier, using the PET technique, Stromswold *et al.* (1996) and Caplan *et al.* (1998) reported selective activation of BA44 when subjects made plausibility judgments about center-embedded relative clauses (15a) compared to right-branching relative clauses (15b):

(15) a. The juice that the child spilled stained the rug.
b. The child spilled the juice that stained the rug.

[8] Similar results have been obtained in other experiments, such as those reported in Tettamanti *et al.* (2002; 2008), which exploited both pseudosentences and non-linguistic symbolic strings within an autonomous learning environment.

To summarize, in this section we pointed out that Broca's Area (BA44 and BA45) is selectively activated in the processing of hierarchical structures. Non-local dependencies (passive constructions and object relative clauses) also involve a special role for BA44.

In the next subsection we will try to relate this selective activation to signatures of complexity during performance, as a preliminary to further discussions of computational complexity in §13.5.

13.4.2 *The complexity effects of hierarchical syntactic processing and non-local dependency formation*

Let us now turn in more detail to syntactic hierarchies and non-local dependencies such as syntactic movement. It appears that as far as these are concerned, an increase of complexity in sentence processing is expressed by an increase of activity in the relevant brain areas. It also appears that working memory is involved in non-local dependencies of the filler-gap type (Wager *et al.* 2003). This means that when constituents become hierarchically deeper and memory load increases (because of increased distance between the filler and the gap and/or because of intervening items), the areas specifically activated would be likely to show stronger metabolic activity or longer metabolic activity and that adjacent areas should be recruited for the more demanding tasks.

All these hypotheses seem to be at least partly borne out. For example, in a pioneering study, Just *et al.* (1996) noted an increase in the volume of neural tissue activation (number of voxels, i.e. volumetric pixels produced by an fMRI imaging system), mainly in Wernicke and Broca's areas, that was proportional to sentence complexity. The sentences they used are given in (16) below:

(16) a. The reporter attacked the senator and admitted the error.
b. The reporter that attacked the senator admitted the error.
c. The reporter that the senator attacked admitted the error.

In (16a), the two conjoined active sentences are on the lower side of the complexity scale. This is because in left-to-right processing, each sentence is uninterrupted and the first DP feeds the required subject positions in both phrases, consistent with canonical SVO order. On the other hand, as discussed in §13.3, (16b) is more complex than (16a). In (16b), the matrix sentence *The reporter admitted the error* is 'interrupted' by the SRC (*the* <u>*reporter*$_i$</u> [*that* $\__i$ *attacked the senator*]). Here too *the reporter* feeds the subject position both in the matrix clause and in the relative clause.[9] This does not happen in (16c), where the relative clause has its own subject, namely, *the*

[9] We have no space to discuss here different interpretations of the relation between the head and the gap in the restrictive relative clause. However both raising and matching analyses require a non-local relation between the head and the gap to be established (see Bianchi 2002 for a review).

senator, which intervenes between the head and the gap in the object position within the ORC (*the reporter$_i$* [*that the senator attacked* $\_{i}$]). As discussed in §13.3, this involves greater processing complexity than with (16a) and (16b) (see Friedmann *et al.* 2009 for a review). In other words, the study of Just *et al.* (1996) tells us that behavioral complexity revealed in psycholinguistic experiments correlates with an activation increase of the areas specifically involved in language processing, and this increase is proportional to hierarchical depth and to the filler-gap distance.[10]

The role of the distance between the filler and the gap must be tested separately from the intervention effects discussed above. For example, looking at scrambling in German, Friederici *et al.* (2006) noted that the greater the distance between the scrambled constituents and their base (canonical) thematic position, the stronger the activation of BA44.[11] Example (17c) below, where both the indirect object and the direct object are scrambled over the subject, is significantly more complex (i.e. there was stronger activation of BA44) than (17b), where only the indirect object is scrambled across the subject; and (17b) was significantly more complex than (17a) (all arguments are in their base position):

(17) a. Heute hat der Opa dem Jungen den Lutscher geschenkt. (S IO DO)
Today has [the grandfather]$_{NOM}$ [to the boy]$_{DAT}$ [the lollipop]$_{ACC}$ given.
b. Heute hat [dem Jungen]$_i$ der Opa _ $_i$ den Lutscher geschenkt. (IO$_i$ S_$_i$ DO)
Today has [to the boy]$_{DAT}$ [the grandfather]$_{NOM}$ [the lollipop]$_{ACC}$ given.
c. Heute hat [dem Jungen]$_i$ [den Lutscher]$_j$ der Opa _$_i$ _$_j$ geschenkt. (IO$_i$ DO$_j$ S _$_i$_$_j$)
Today has [to the boy]$_{DAT}$ [the lollipop]$_{ACC}$ [the grandfather]$_{NOM}$ given.
'Today, the grandfather has given the lollipop to the boy.'

According to Grodzinsky (2000), BA45 also seems to play an important role during processing of certain kinds of non-local dependencies. The evidence is based both on aphasic patients with selective lesions and fMRI experiments (Santi and Grodzinsky 2007).[12]

Makuuchi *et al.* (2009), in a 2-way factorial design study, compared the filler-gap distance (short vs. long) component with the level of embedding (see sentences 18a–b, where the first involves embedding and the second does not):

[10] This study however does not help us to understand which component of the complexity function, as we presented it in §13.3, is involved, because the distance between the filler and the gap (that is, Gibson's 1998 memory-load cost) and the nature of the item that intervenes between the filler and the gap (that is, what we call Feature-based Integration Cost 8) are non-identical in (16b) and (16c).

[11] Ben-Shachar *et al.* (2004), testing topicalization vs. dative-shift in Hebrew, reported a more articulated activation of BA44/45 and BA6/9. We do not have space to discuss this matter here. It is simply worth stressing that scrambling and topicalization might involve different neural pattern activation. This is surely true for shifting operations like dative-shift (Larson 1988; Ben-Shachar *et al.* 2004).

[12] Santi and Grodzinsky (2010) found an activation on BA45 during syntactic movement processing, but not when just pronominal binding was involved.

(18) Peter wusste, dass (Peter knew that) . . .

a. [Maria$_1$, [[die Hans]$_2$, [[der gut]$_3$ aussah$_3$] liebte$_2$] Johann geküsst hatte$_1$].
Maria who Hans who was good looking loved Johann kissed. [literal]
'Maria who loved Hans who was good looking kissed Johann.'

b. Achim$_1$ den großen Mann gestern am späten Abend gesehen hatte$_1$.
Achim the tall man yesterday at late night saw. [literal]
'Achim saw the tall man yesterday late at night.'

The observed activation patterns suggest a contrast between BA44, which is sensitive to structural embedding, and the dorsal portion of both BA44 and BA45 (i.e. the Inferior Frontal Gyrus), which is involved in memory-demanding tasks requiring movement. Using an imaging technique that allows us to draw probabilistic connections between areas by looking at movement of water molecules within the brain tissues (diffusion tensor imaging), Makuuchi *et al.* (2009) also identified a strong connection between the dorsal portion of the two regions BA44 and BA45. The increase of activation is present both for nesting and non-local dependencies. In other words, these experiments were only partly able to disentangle the role of memory from the role of hierarchy.

To conclude, in this section we showed how specific brain regions that are active during hierarchical syntactic processing and non-local dependency formation become more active when hierarchy gets deeper, as in relative clause embedding, and when dependencies require extra working memory, as is the case with longer dependencies containing constituents that structurally intervene between the filler and the gap. We can now go back to our computational complexity model to draw some conclusions from what we learned.

13.5 Possible and impossible rules and automata

This section concludes our discussion by relating the computational complexity findings of §13.3 to the neurophysiological evidence we presented in §13.4. What we discovered is that the distinction between possible and impossible rules, as investigated by Musso *et al.* (2003), lends itself to modeling in terms of FSAs and PDAs. (19–20) repeat their crucial sentences, while (21–22) formalize their results in terms of automata:

(19) Possible rules

a. **La pera** è mangiata **da Paolo** (passivization)
the pear is eaten by Paolo

b. Pia dice **che** Paolo mangia la pera (complementation)
Pia says that Paolo eats the pear

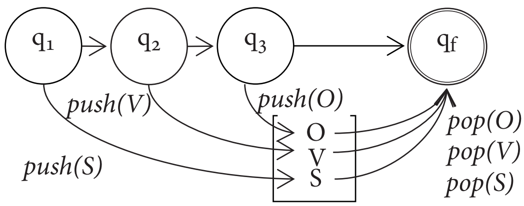
(20) Impossible rules

a. Paolo mangia la **no** pera (negation)
Paolo eats the no pear

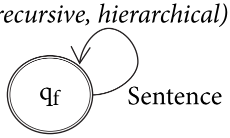
b. Un**a** bambino mangia una per**a** (agreement)
$a_{fem.sing}$ child eats a $pear_{fem.sing}$

(21) Possible rules

a. Given a sentence, passivize it by inverting the subject and the object:
(non-recursive, hierarchical)

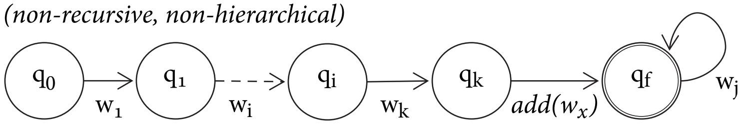


b. Expand a sentence with another sentence by complementation
(recursive, hierarchical)

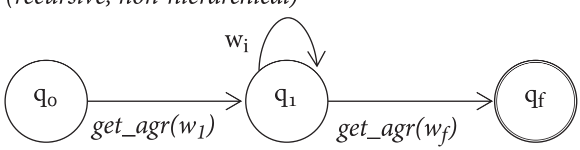


(22) Impossible rules

a. Insert a word w_x at k^{th} position (requires $k+1$ states);
(non-recursive, non-hierarchical)



b. The first, w_1, and the last element, w_f, in the string should agree
(recursive, non-hierarchical)



We observe that only passivization, which involves a non-local dependency, requires a device that is as powerful as PDA. Complementation simply requires FSA power. This is an important point, since the activation of the areas (BA44 and BA45) discussed in Musso *et al.* (2003) seems sensitive to hierarchy, but not to recursion or to memory demand.[13]

13.6 Conclusion

In this chapter we investigated the relationship between complexity, as defined by the Chomsky hierarchy and its associated hierarchy of automata, and the relative complexity of operations in the brain. What we found is that two main components are needed in order to correctly characterize a complexity function that is computationally sound (§13.2), psycholinguistically tenable (§13.3) and neurophysiologically testable (§13.4). These components are the hierarchical embedding (time complexity) and the memory demand expressed in terms of intervening features within a filler-gap dependency (space complexity). We hypothesize that our results, which are based on studies of a few familiar languages, will generalize to human language in general. Whether this is or is not the case, as well as the question of whether all languages avail themselves of the same neurologically-instantiated mechanisms of the same degree of complexity, are matters for future research.

[13] Because of Makuuchi *et al.*'s (2009) experimental results, we should expect a stronger activation of the dorsal portion of BA44 and BA45 proportional to the complexity of the non-local dependency, §13.4.2. Since no difference in pattern activation is revealed by Musso *et al.* (2003) in (21a) vs. (21b), we conclude that the feature-based integration cost, discussed in (8), is rather low in the passive construction task exemplified by (21a).

14

Looking for a 'Gold Standard' to measure language complexity: what psycholinguistics and neurolinguistics can (and cannot) offer to formal linguistics

LISE MENN AND CECILY JILL DUFFIELD

14.1 Introduction

14.1.1 *Why do we need a gold standard and why don't we have one yet?*

A complexity metric for languages needs a 'gold standard'—that is, an objective way to validate relative language complexity that is theory-independent and anchored in the human mind. Without an anchor in human behavior, any metric of language complexity remains an abstraction from actual user effort, and may not correspond properly to that effort.

To begin, any proposed formal whole-language complexity measure is several steps removed from human behavior. The first 'remove' between a language complexity measure and the individual comes from the fact that a complexity measure can look only at the language-out-there ('E-language', Saussurean *langue*). But effort is internal to the individual, and it is our individual mental grammars ('I-language') which are acquired and which—collectively—determine the shape of the external language and how it will change over time. Therefore, the first reason that a complexity metric based on the (E-) language as a whole needs behavioral validation is that such a metric can, at best, be only an approximation to the aspects of complexity that are important for understanding how language develops in the individual and over historic time.

The second reason that proposed complexity metrics need validation by looking at how people actually behave is that even if we had a formal 'I-language' grammar for an individual, a structure-based measure of its complexity might not correspond to what is difficult for that person to produce and understand. We cannot know *a priori* what is simpler for our minds; we have to look at what is harder or easier for people to learn, produce, and comprehend in spoken, signed and written language, and try to figure out, from those empirical measures, what might be going on inside an individual's head.

Consider an example of how our minds adapt as we use (written) language. The French spelling system is highly conservative, in the sense that it reflects the history of the language. Many words that are (now) pronounced as homonyms retain spelling differences that the learner must memorize, such as *tous* ('all') and *toux* ('cough'), both pronounced /tu/. Learning the different spellings increases complexity for the beginning reader. However, after the words are learned, the separate spellings decrease ambiguity and should therefore speed decoding. In other words, the separate spellings decrease complexity for the skilled user. More generally, users become specialized for processing their language(s). A property that makes a language hard to learn may make that language's information easy to process once one has learned it. No presently available description, formal or information-theoretic (Juola 1998, 2008), deals with this kind of variation.

Third, even looking at the individual does not break the complexity problem down far enough. Within the individual, we have to consider modality (speaking, hearing, writing, reading, signing). Reading and writing obviously permit longer processing times than speaking and signing, and increase the ability to encode and decode embeddings. On the other hand, intonation, tempo, and pauses give the hearer far more structural information than punctuation gives the reader.

There are also significant speaker/hearer asymmetries: the speaker has to deal with the problem of getting a scenario into a linear string of words, which requires, among other things, choosing a 'starting point' (see Bock, Irwin, and Davidson 2004 and citations therein).

Context also affects the difficulty of processing an utterance. Listeners make use of real-world contexts to aid in disambiguating garden-path sentences during comprehension (Spivey, Tanenhaus, Eberhard, and Sedivy 2002), and real-world contexts influence speakers to encode information above and beyond what is necessary for referential success in the discourse context (Wardlow and Ferreira 2008).

In summary, while a complexity measure that is independent of a user's state of knowledge provides a source for hypotheses to test, proposed whole-language complexity measures remain psychologically problematic because they do not deal with the individual, they do not change with the individual's experience, and they do not deal with modality differences, speaker/hearer asymmetries, or the effects of context.

14.1.2 *Basic constraints on complexity measures*

Complexity measures for a language as a whole should reflect the distribution of factors that contribute to the difficulties that hearers and speakers face when they are processing utterances. How often and under what conditions do users of a particular language experience processing problems? Counting the frequency of each potentially difficult structure is not enough; what matters is the frequency with which those structures occur in contexts that actually impose a processing burden on the speaker or hearer. For example, computational linguists can tell us how often center-embedded and multiply center-embedded structures occur in various corpora. But center-embedded clauses can be relatively easy to understand when the semantic relationships between the subjects and the verbs are obvious to the hearer, e.g. *The report the senator the car hit wrote impressed the president* compared to *The rat the cat the dog chased killed ate the malt.* And perhaps there are trade-offs—for example, there might be circumstances under which center embeddings are easier to produce even though they are harder to understand.

So valid complexity measures must deal with the presence of presumably difficult structures and forms and with how frequently speakers and hearers encounter them in various contexts of use. These measures must further deal with whether, in those contexts, those forms really pose difficulties for speakers and for hearers, for writers and for readers.

Since there cannot be a single measure of the complexity of an utterance, a reasonable goal for psycholinguistics and neurolinguistics would be an empirically-supported list of the principal linguistic factors that increase complexity for speakers and for hearers (and readers and writers) with various levels of expertise. Psycholinguistic and neurolinguistic research should also provide a sense of how these features interact with non-linguistic cognitive factors, like working memory and knowledge of prior or current context.

14.1.3 *Empirical methods for testing relative utterance complexity*

Figuring out how to resolve a particular claim about the relative complexity of an utterance for speakers or hearers using available methods can be challenging. The kinds of psycholinguistic data that show A to be simpler to understand or produce than B (when other factors are held constant) include the following: A is processed faster than B; A is processed more accurately than B; a person can do other tasks better while they are doing A than while they are doing B; and people are more likely to err in the direction of saying A when they intend to say B. Experimental neurolinguistic methods like ERP (event-related potential) can also show whether less brain effort is needed to produce or comprehend A than B, and clinical neurolinguistic studies analyze the nature of aphasic errors.

In comprehension studies, measures of processing effort include reading times and reaction times in decision tasks (with longer times assumed to correlate with greater effort); eye tracking (where location and duration of gaze reflect material being processed and the effort it takes to process it); and ERP studies. ERP measures the signal-averaged EEG responses to linguistic events (as well as other cognitive and motor events). Greater (positive or negative) ERP values are assumed to reflect higher processing effort (e.g. in processing novel vs. familiar metaphors; Lai, Curran, and Menn 2009). In production research, the most common measure of processing effort is error rate: a higher proportion of errors in one condition than in another reflects increased processing load. Other methods used in production studies include measuring onset latencies (the time it takes to initiate an utterance), looking at word-order preferences, and tracking eye movements during speaking.

14.1.4 *Complexity and competition*

Let us turn now to another contributor to complexity which illustrates the importance of competition in language processing: 'garden-path' sentences. Garden-path sentences are those whose structures are temporarily ambiguous, and whose words pull the reader or listener towards a parse that cannot be completed. Tom Bever's notorious sentence *The horse raced past the barn fell* (Bever 1970) must be parsed as having a reduced relative clause modifying the subject, i.e. as 'The horse that was raced past the barn fell.'

As is well-known, getting pulled towards the wrong parse does not depend only on the presence of a (temporary) structural ambiguity (at least for adults; see Trueswell, Sekerina, Hill, and Logrip 1999 for a review of adults' and children's processing of garden path sentences). The sentence *The ball rolled through the end zone scored* is structurally equivalent to Bever's monster, but it is much easier to understand. We seem very quick to take *The horse raced past*... as the subject and main verb of a single clause describing an agent-subject intransitive ('unergative') action, but not as quick to decide whether *The ball rolled through*... is the subject and part of the predicate of a single main clause or the subject of a main clause followed by the beginning of a reduced relative clause. This seems to be because of real-world likelihoods and the way humans conceive of them.

However, years of experimental work have shown that having to choose between potential parses slows down readers' processing even when they do not consciously experience having any problem understanding a sentence. In fact, one of the best-understood contributions to the difficulty of comprehending an utterance is the amount of competition among possible potential parses (Trueswell and Kim 1998). This competition is modulated both by the relative frequencies of the possible parses and by the plausibility of the interpretations that those parses yield. (See Thothathiri *et al.* 2012 for further analysis and discussion.)

More generally, a word or a structure in a context where it is relatively predictable is easier to comprehend than in a construction where it is unexpected. Predictability helps comprehension because there is intense, automatic, subconscious competition among potential words, parsings, and semantic interpretations. Decoding a message is harder when the correct parsing is the less probable one, as in *The horse raced past the barn fell.* And at the level of word comprehension, "speech processing involves automatic simultaneous activation of multiple candidate words that are fully or partially consistent with the input, as well as competition between these candidates, with the competition modulated by probabilistically weighted segmentation and selection procedures" (Cutler 2012: 191).

Competition among possible ways to proceed is not a problem only for comprehension. It is also rampant in sentence production, and it is a major contributor to the complexity of producing utterances. Other things being equal, more frequent and more predictable items and structures are more accessible to the speaker. Therefore, highly predictable items have an easier time winning the competition for what to say next. Conversely, conveying a message is more difficult when it requires using a less likely word or structure. So understanding and evaluating complexity requires facing the issue of how frequency and usage facts interact with more traditionally-recognized aspects of grammar.

14.1.5 *Psycholinguistic models of representation and processing*

If the validation of language complexity measures must be based on complexity-for-human-language-processing (including learning and the effects of non-linguistic factors), then creating such measures depends on creating sufficiently good psycholinguistic models of language processing. Current psycholinguistic models have yet to deal with variation in the user's mental state. Furthermore, despite budding research in the links between comprehension and production (see Gennari and MacDonald 2009; MacDonald and Thornton 2009), no single processing model yet encompasses the variation across speakers and hearers. This chapter examines evidence concerning language processing, both psycholinguistic and neurolinguistic, bearing on the question of complexity, and the implications of that evidence for modeling language processing and language representation. (Evidence from language learning is beyond the scope of this chapter.)

We conclude by proposing a psycholinguistic model of language representation and processing that might provide a heuristic basis for integrating many of the factors that contribute to the complexity of an utterance. If it can be sufficiently elaborated, the model might be the basis for a computationally explicit language processing model which would be a working approximation to the behavioral gold standard that theoreticians need.

Our proposed psycholinguistic model provides a way to handle the interaction between hierarchical structure, sequential structure, and frequency. In the model, usage information is distinguishable from information about hierarchical structure of a sentence, but usage creates syntagmatic structural units, which also play a role in sentence processing. Some of these may, over time, be re-analyzed as elements in the hierarchical structure. Hopefully, our model thus responds to some of the concerns of both usage-based and formalist theorists. It is presently only a spoken language production model, but its architecture is not modality-specific, so it should permit expansion to comprehension and to written language. While the model does not yet handle non-linguistic factors that also affect complexity (working memory, non-linguistic context), it should be hospitable to elaborations that could do so.

14.2 Problematic data: disparities between formal complexity and processing difficulty

14.2.1 *Paradigmatic competition increases complexity*

Data from people with aphasia were classically taken to support general (cross-linguistic) notions of markedness (for phonology, Jakobson 1956; for morphology, LaPointe 1985). However, cross-linguistic data that have become available in recent decades show that word frequency and co-occurrence frequency can override markedness. So, while some aphasic speakers (notably, people with agrammatism) tend to use the singular form for plural nouns, a favorite example showing the importance of frequency of co-occurrence in morphology is Dressler (1991). In that study a Breton speaker with fluent aphasia tended to name pictures or examples of a single object using the plural form if the object itself was most frequently found in quantity (leaves, potatoes) and, even more dramatically, using the *dual* if the object was usually found in pairs (eyes, hands). There are parallel effects in experimental studies: see Biedermann *et al.* (2012) for English speakers with aphasia and Kuperman *et al.* (2009) for speakers without brain damage.

The relative frequency of a form in its inflectional paradigm is important because of its effect on resolving the covert competition among the members of the paradigm. Within-paradigm competition becomes evident from the way speakers with aphasia often flounder among possible words and inflected forms. For example, aphasic German speakers may struggle to find the right forms of the definite article (six forms distributed somewhat chaotically among four cases and four gender/number categories); adjective gender agreement is a problem for some aphasic French speakers; English agrammatic speakers often have problems with past tense and 3rd person singular suffixes. (See Bates, Wulfeck, and MacWhinney 1991 for German, also Italian; Bastiaanse *et al.* 2003 for case errors in Dutch; Menn and Obler 1990 for English). In agrammatic speakers of Slavic languages, nominative or accusative forms often replace required instrumental, locative or genitive forms (e.g. Marková and Cséfalvay 2010), even when the speaker gives evidence of attempting to place the words in construction with a governing verb or preposition. In summary, forms with greater frequency have an advantage in these competitions, whether they are correct or not.

Where there is less paradigmatic competition, aphasic speakers do better. For example, Japanese people with aphasia, spared the labor of gender and number marking, appear to have little difficulty choosing correctly among the three commonest verb aspectual forms (non-past finite *-masu*, past finite *-ta*, non-finite/conjunctive *-te*) (Sasanuma, Kamio, and Kubota 1990). Both structural and statistical factors appear to affect how hard it is for aphasic speakers to resolve competition in favor of the correct forms (Nicol, Jakubowicz, and Goldblum 1996). These findings, then, would concur with the structure-based prediction that a more elaborate paradigm is more complex, but they add to it the usage-based finding that the more difficult inflected forms in a paradigm are those which are less frequent. Furthermore, a key aspect of frequency is not the overall frequency of, e.g. the dative plural, but the frequency of the particular inflected form with respect to the other inflected forms of the same stem.

14.2.2 *Hierarchical structure affects but does not fully determine complexity*

While additional structure can result in additional complexity for the language user, there is no one-to-one relationship between structure and processing difficulty. We have already mentioned the fact that predictability—a usage factor—reduces processing difficulty, but even when predictability is held constant, some kinds of differences in processing difficulty are hard to explain in structural terms. For example, as is well-known, speakers find it harder to comprehend sentences with several center-embedded relative clauses than to comprehend very similar sentences with the relative clauses embedded later (Karlsson 2007). Compare (1) and (2):

(1) The report the senator the car hit wrote impressed the president.

(2) The car hit the senator that wrote the report that impressed the president.

Central to the discussion of structural accounts of complexity are constraints on which elements can appear outside of (or 'be extracted from') certain structurally defined environments. These are typically referred to as 'island constraints' (Ross 1967). Examples include the Complex NP Constraint and the WH-Island Constraint, both of which have been encompassed by the Subjacency Principle (Chomsky 1973). Constraints such as these are proposed to rule as ungrammatical such utterances as those in examples (3) and (4), taken from Hofmeister and Sag (2010):

(3) ?It was a new company that Simon spread the rumor that they started.

(4) ?What did the attempt to find end in failure?

But acceptability judgments vary (author LM finds 1 acceptable, 2 unacceptable; author CJD is more lenient). Furthermore, exceptions to these constraints have been demonstrated across languages. Hofmeister and Sag (2010) show that some of these exceptions can be explained by non-structural features. In particular, properties of the extracted item (*which*-phrases vs. bare *wh*-items) and of the structure from which it is extracted (definite vs. indefinite complex noun phrases) affect how easily these violations are processed, and the processing measures correlate with off-line judgments of the acceptability of these structures.

Hofmeister and Sag propose a processing account of island constraints. In their analysis, cognitive constraints, rather than structural properties, render certain structures more or less acceptable, and more or less complex for the language user. Similar structures can be more or less difficult to process depending on non-structural properties (e.g. definiteness). If these properties make it hard to process the material intervening between the filler and the gap, it will be harder to retain the memory representation of the filler item and reconcile it with the gap. For example, a bare WH-item extracted from a definite NP (in 5) is shown to be more difficult to process than an identical structure in which a WH-phrase is extracted from an indefinite NP (in 6):

(5) I saw **who** Emma doubted **the report** that we had captured in the nationwide FBI manhunt.

(6) I saw **which convict** Emma doubted **a report** that we had captured in the nationwide FBI manhunt (Hofmeister and Sag 2010: 388).

Some other non-structural factors can affect the processing of filler-gap dependencies in comprehension tasks. These include the introduction of discourse-new information (Gibson 1998), the 'givenness' status of referents in embedded clauses (Warren and Gibson 2002), and the similarity among noun phrase types (Gordon, Hendrick, and Johnson 2004; Gordon *et al.* 2006).

Such findings are problematic for complexity metrics defined strictly in terms of hierarchical structure. As Hofmeister and Sag (2010) note, when non-structural factors render structural violations more-or-less acceptable, a formal theory that does not acknowledge processing constraints is forced to deal with such utterances in an ad hoc fashion, ruling them 'in' or 'out' instead of variable, and in some cases, setting aside counterexamples as peripheral to the issues under discussion.

14.2.3 *Structures may be processed differently across production and comprehension*

Modality-specific effects pose a challenge to the assumption that grammar is amodal or that the grammar used in production is a proper subset of the grammar used in comprehension. Of course comprehension and production grammars must interact. MacDonald's Production-Distribution-Comprehension model (Gennari and

MacDonald 2009; MacDonald and Thornton 2009), for example, claims that pressures on the production system determine the distribution of linguistic forms available as input for the listener or reader. But production and comprehension grammars need not necessarily be identical. Whether or not they are remains an open question, although one that is discomfiting to most linguists. Even Sag and Wasow specifically posit an amodal grammar, which seems to mean that all production–comprehension asymmetries would have to be accounted for by online processing demands.

One recent study demonstrating an asymmetry in comprehension and production is Francis and Michaelis-Cummings's (2012) investigation of features contributing to the distribution of extraposed relative clauses. Francis and Michaelis-Cummings used a preference task and two anagram production tasks (participants were shown three constituents—subject NP, relative clause and verb phrase—on a screen and had to create a sentence using them). One version of the anagram production task required remembering the constituents, the other did not. Francis and Michaelis-Cummings investigated whether definiteness of subject NP, relative clause length, and VP length would affect language users' preferences for relative clause extraposition. In the preference task, which involved only making judgments, there was no main effect of VP length. In other words, participants judged (7), with a long VP and extraposed relative clause to be just as acceptable as (8), which has a short VP.

(7) Some research has been conducted fairly recently that refutes the existing theories.

(8) Some research was conducted that refutes the existing theories.

But in the anagram sentence-creation tasks, VP length was a significant predictor, with short VPs predicting relative clause extraposition (cf. Heavy Noun Phrase Shift, Ross 1967). In other words, participants would be more likely to produce (7) than (8). Relative clause length was significant in the judgment task and the memory load production task (for the effect of definiteness, see Francis and Michaelis-Cummings 2012).

Why should language users treat VP length differently in different tasks? We suggest that phrase length may impose different kinds of processing loads for the speaker and the comprehender. During production of an utterance, speakers must be able to access semantic and syntactic representations very quickly in order to produce a fluent utterance. As MacDonald and Thornton (2009) note, shorter phrases are generally more accessible (that is, the speaker can assemble them more rapidly), and they are therefore more likely to be produced earlier in the sentence than longer phrases, unless their function makes that ordering undesirable. This explanation is consistent with the preference for ordering short VPs before relative clauses in the Francis and Michaelis-Cummings study.

During the comprehension of an utterance, shorter interruptions should also reduce working memory load, leading to a preference for postponing longer phrases. And according to MacDonald's Production-Distribution-Comprehension model, patterns in production can lead to preferences for similar patterns in comprehension, at least when interpreting structural ambiguities. However, the conditions under which the expected preference for postponing longer phrases will emerge are not clear. In the Francis and Michaelis-Cummings study, the preference/judgment task presented no ambiguities—the relative clause clearly modified the subject NP, even when following a longer intervening VP—and showed no preference for shorter interrupting phrases.

A more direct investigation of the differences between production and comprehension comes from Tanner's (2012) revisiting of Bock and Cutting's clause-bounding effects in subject-verb agreement. Bock and Cutting (1992) had investigated the effect of local intervening nouns on the speaker's successful production of correct verb number. Measuring error rates in number agreement on the verb, they demonstrated that a local noun is less likely to interfere with subject-verb agreement if it is embedded in a relative clause than if it is embedded in a prepositional phrase. For example, a speaker is more likely to erroneously produce a plural verb (despite the singular subject) in (9) than in (10):

(9) The advisor [$_{PP}$ for the chemistry students] . . .

(10) The advisor [$_{RC}$ who directed the students] . . .

This pattern is surprising if we simply assume that additional structure should make processing the grammatical dependency more difficult. Bock and Cutting propose that rather than increasing the complexity for the speaker, the clausal structure actually helps to encapsulate the information. Simply put, what happens inside the clause stays inside the clause. This 'clause bounding effect' keeps the local noun from interfering with the speaker's coding of number on the verb in the main clause.

But the clause bounding effect does not appear to hold in comprehension. Tanner (2012) used various methods (ERP and self-paced reading tasks) to demonstrate that when reading sentences containing agreement errors, language users are just as likely to be distracted by an intervening noun embedded in a clausal modifier as one embedded in a phrasal modifier. In other words, having an intervening plural noun in either a prepositional phrase or a relative clause modifier of a singular subject noun phrase appears to keep readers from noticing an erroneous plural verb in the main clause.

The difference in results may be due to task demands. Memory tasks demonstrate that language users have stronger retention for the meaning than of the surface forms of recently perceived utterances (Sachs 1967). However, surface structure can be retained if language users know that they will be immediately tested for verbatim information (Johnson-Laird and Stevenson 1970). And verbatim recall is required when a participant in an experiment must repeat a subject preamble and create a sentence with it. While recall measures are much less sensitive to on-line comprehension processes than ERP and self-paced reading, it may be that participants in comprehension tasks do not pay as much attention to surface forms as they do in production experiments.

It will take much work to determine whether these differences should be accounted for in terms of task demands, in terms of modality-specific grammars for speakers and comprehenders, or both. Understanding the relationship between production and comprehension, and how complexity is to be defined in each realm, will require looking at a wider range of tasks, analyzing the results in terms of a (sufficiently detailed) psycholinguistic model, and deciding what belongs in a grammar and what should be relegated to processing (if, indeed, grammar exists as more than an abstraction from or an idealization of processing).

14.3 Frequency, co-occurrence frequency, and implications for representation

If the complexity of a structure for a user depends partly on the structure's relative frequency, then the frequency of structures must be represented in language users' minds. What is this representation of frequency like, and how does a user's experience create it?

14.3.1 *Lexical frequency and its interaction with lexical and semantic category*

Word frequency is the strongest predictor of how rapidly a word can be read aloud (holding the number of syllables in the word constant), and of how rapidly a picturable object can be named (Oldfield and Wingfield 1965; Carroll and White 1973). Frequency is never the sole determinant, however. For example, normal speakers read concrete nouns aloud or judge them to be words relatively faster than equally frequent abstract nouns, and faster than words in any other syntactic category. People with aphasia may have category-specific difficulties with nouns, with verbs, with functors, or with semantic subcategories of nouns (e.g. inanimate objects, numbers, colors, body parts, Hillis and Caramazza 1991).

The psychological and neurological bases for these 'other factors' are matters of intense research and debate. But the basis for the frequency effect is so uncontroversial that it is rarely discussed: the internal representation of a word is somehow strengthened each time it is heard or spoken. This does not mean that frequency effects require counting, any more than muscles have to be able to count the repetitions of an exercise in order to get stronger, but they do require that a word's representation be the kind of cognitive object that can be incrementally strengthened each time it is used. Neural net models are the best available conceptual approximation to understanding how the brain works (even though they are vastly simpler than the brain), because the brain is assuredly a network, and such models have many ways of automatically strengthening representations with experience (Smolensky and Legendre 2006).

14.3.2 *Co-occurrence probability, processing complexity, and representation*

How do speakers acquire and store information about the relative frequencies of forms in an inflectional paradigm?

14.3.2.1 *Representing inflectional linkages* We have seen that it is easier to process inflections when they are attached to words on which they frequently appear than when they are attached to words on which they occur less frequently. But inflectional morphemes and lexical stems are different kinds of psycholinguistic objects, as shown by normal speakers' slips of the tongue. In 'stranding errors', the lexical items move around, leaving inflections (and often, functors) in the places where the speaker intended them, e.g. *churchES in our minister* (for *ministerS in our church*); *A sentence is a bad thing to end a preposition WITH* (Fromkin 1993, Boulder speech error corpus). This poses a puzzle, namely, how can we represent the tightness of the linkage between a lexical item and its inflection if they have different sources in the production mechanism? Furthermore, if this tightness depends on our experience of hearing a particular inflected form of a stem more frequently than other forms of that stem, how can we represent the development of this affinity over the learner's lifetime?

In a model that uses only hierarchical structures, an inflected word must be represented either as a composite assembled from stem and affix or as an amalgamated whole. A learner would have to switch from one kind of representation of an inflected form to the other, or from having one of them one to having both. When and how would this postulated switch happen? (An extensive psycholinguistic/neurolinguistic literature debates these matters, e.g. Schreuder *et al.* 1990.) Hierarchical representations (and the classical production model, which augments hierarchical structure only with lexical frames) are poorly suited to representing syntagmatic relationships like those between stems and affixes.

But network models can provide representations of the syntagmatic linkages between successive morphemes in an inflected word (or in a construction), as well as gradient-strength representations of these linkages and of lexical items. Furthermore, as Bybee (2007) argues, networks can model paradigmatic connections across semantically and formally similar sets of words (or sets of constructions). Forming these cross-connections is the way networks instantiate the young speaker's growing recognition of the morphemes and larger structural units of the ambient language. If an inflected form is so common that children learn it first (or if it is irregular), its decomposition develops later, as they find correspondences between that form and the inflected forms of other words. Finding the decomposition does not mean abandoning the holistic representation, but rather, augmenting it.

14.3.2.2 *Representing syntagmatic structures* Because network representations can be both syntagmatic and paradigmatic, linking parts of a form to structurally-similar parts of other forms removes the need to postulate separate unanalyzed and analyzed representations for inflected forms, and for longer syntagmatic structures as well.

Idiom blends provide more evidence that the 'stored or assembled' dichotomy is counterproductive. Idioms do not lose their internal syntactic structure just because they are semantically opaque (although they may lose syntactic flexibility; Culicover 1999). Consider two observed speech errors: 'Did someone fly off the deep end?' (a blend of 'fly off the handle' and 'Jump in the deep end', metaphors respectively for losing control of one's temper and of one's actions), and the reprimand 'How do you sleep with yourself?' (a blend of 'How do you sleep at night?' and 'How do you live with yourself?'). Both of these blend errors show speakers producing strings that respect the internal structure of each component idiom, rather than producing a surface string that combines them into a string of gibberish (e.g. 'Did someone fly in deep the handle end?' or 'How do you sleep at with yourself?,' cf. Cutting and Bock's 1997 experiments eliciting idiom blends showing that speakers respect the internal structure of idioms). In summary, a good way to think about conventionalized composite linguistic objects is as composites with particularly strong internal links. On this view, semantic idiomaticity and syntactic analyzability are separable gradient properties of an expression, rather than being diametric opposites.

14.3.2.3 *Sequential co-occurrence probabilities create syntagmatic structures* Learning sequential co-occurrence probabilities is like learning word frequencies: we learn from instances and encode their relative frequencies in the strengths of links between successive structures, between structures and the lexical items that occur in them, and between successive lexical items. This gives our knowledge of language a syntagmatic structure in addition to its hierarchical and paradigmatic structures. For example, we learn that locative expressions are likely to occur in particular places in clauses (*Throw mama a kiss from the train*, not *Throw mama from the train a kiss)*, and are likely to either be adverbs or prepositional phrases. We further learn that particular adverbs and particular types of prepositional phrases are likely to be used in specifying locations (*under the bridge* is more likely than *under the tracks*), and we draw on this knowledge during comprehension and production.

Purely sequential strings can become strongly linked, forming collocations regardless of whether they sit neatly under a hierarchical-structure node (like a nominal stem and its inflectional affixes) or not (like subject-auxiliary contractions). The 'strong internal links' notion of collocations allows us to deal with the fact that frequent collocations are simpler for aphasic speakers to produce or repeat. This is clinical common knowledge. For example, scoring systems for an aphasic speaker's discourse (e.g. for the Boston Diagnostic Aphasia Examination (BDAE), Goodglass, Kaplan, and Barresi 2001) instruct the clinician to set aside 'idioms and frequent expressions' (*I'm doing okay*) when rating the fluency of the speaker's speech production. Furthermore, repetition and comprehension tests for aphasic speakers are built (although not systematically) around the recognition that highly probable sequences of words are easier to process than unusual ones with the same syntactic structures. Compare, for example, the 'high-probability' *I got home from work* with 'low-probability' *The spy fled to Greece* (both from the BDAE). Although the differential difficulty of these particular sentences is compounded by the great differences in the frequencies of their component words, the effect of speakers' context-based expectations on word recognition is well-known, as for example, in studies of the phoneme restoration effect (Warren 1970) and category boundary shift (Ganong 1980).

14.3.2.4 *Simultaneous co-occurrence probabilities affect processing complexity*

Simultaneous co-occurrence frequency is also learned from instances, and language users automatically compute it across the whole range of linguistic levels from phonetics to pragmatics. Take an example from sociophonetics as an illustration: English users learn information that links phonetics with grammatical category. Although /ð/ is the most frequently occurring consonant in English, in word-initial position it is restricted to function words (*the*, *these, this, thine*. . .). As a result, it is difficult for native English speakers to use #/ð/ at the start of a content word. So the Greek word for stuffed grape leaves, /ðolma/, is difficult and the plural /ðolmaðes/ is even harder.

The /ð/ example is an absolute, because #/ð/ never occurs in English content words, but probabilistic learning also links levels. For example, the likelihood that a verb occurs in a particular syntactic frame affects how easy it is to process for speakers with aphasia. Consider verbs that can occur in several different syntactic frames, such as *walk*, which occurs in active, passive, and agent-subject intransitive ('unergative') frames, or *shrink*, which occurs in transitive, passive, and undergoer-subject ('unaccusative') frames. Speakers with aphasia can understand such a verb relatively more easily when it is in the frame in which it most frequently occurs. In a plausibility judgment task (Gahl *et al.* 2003), the passives of English verbs that are more frequently used in the passive than in the transitive active (*injure, arrest, elect*) were easier for people with aphasia to understand than passives of verbs that are most frequently used in the active voice. Also, unaccusative sentences containing verbs which are most frequently used in unaccusative frames, such as *burst* and *float*, were slightly easier to understand than transitive active sentences made with the same verbs.

14.3.3 *Representing constructions and collocations*

The information that a word belongs to a collocation or a construction must be represented some how in our minds, because when a word occurs in a collocation or construction where it is highly probable, it is easier to process. How is it possible for our minds to represent syntagmatic structures as well as hierarchal structures? Before discussing the representation of syntagmatic structures, we need to define constructions and collocations more explicitly. While collocations and constructions often overlap, they are distinct types.

Goldberg defines a construction as a conventionalized, learned form-function pairing. Constructions exist at all levels of abstraction, from the unfilled 'Subj-V-Obj_1-Obj_2' ditransitive construction to specific lexical items (words being also 'learned pairings of form and function'). While most utterances are not constructions, any utterance is comprised of multiple overlays of constructions (Goldberg 2006). For example, the utterance, 'Good grief, aren't these well-fed cows!' is a spontaneous, creative utterance comprised of an idiomatic expression (*good grief*), the copular construction, the subject-aux inversion construction, a negative-copula contraction, and various phrasal (VP, NP) and lexical constructions. The recognition that a single element of an utterance may simultaneously be part of any number of the constructions within a single utterance (e.g. the *aren't* as part of the subject-auxiliary inversion, the copular, and the contraction constructions) allows us to discuss processing effects at various levels. In terms of a formal grammar, consider the multiple inheritance hierarchy of Sign-Based Construction Grammar (Sag 2012) that defines the relations between networks of constructions. In such a framework, any construction can be a subtype of multiple supertypes, inheriting their features. A single utterance can comprise multiple constructions, inheriting processing advantages or disadvantages from the features of each of those constructions.

In a gradient or network view of grammar, a construction will be relatively 'strong' if it is frequent and/or if it has a well-defined function, but it will be weak or marginal if it is rare or hard to differentiate from similar structures. (Reading through the *Oxford English Dictionary*'s enormous entry for *of* gives an intuitive sense of what 'strength' means. Their discussion says "All the existing uses of *of* are derivative; many so remote as to retain no trace of the original sense, and so weakened as to be in themselves the expression of relatively indefinable syntactic relationships . . . From its original sense, *of* was naturally used in the expression of the notions of removal, separation, privation, derivation, origin or source, starting-point, spring of action, cause, agent, instrument, material, and other senses, which involve the notion of 'taking, coming, arising, or resulting from', ..."

We use 'collocation' to refer to a set of words or words and grammatical morphemes that are used together in sequence so frequently that they behave in many ways like a unit, regardless of whether they hang from a single node in a hierarchical structure. Idioms, formula frames (*I'd like to introduce*...), and contractions are prime examples. 'Collocation' is also a category with fuzzy boundaries. However, collocations cannot be dismissed as 'marginal' to linguistics, because central synchronic and diachronic linguistic phenomena occur only within collocations (cf. Bybee 2007). Some examples in Indo-European languages are contraction, clitic formation, and affix formation. While many listed collocation types are constructions or instantiations of constructions, some of the fluent and intact-sounding runs of speech found in aphasia such as *in the attic* and *I'm gonna hafta run outside* (from a speaker with Wernicke's aphasia recounting 'Cinderella', Berndt 2001) do not appear to be linked with any special conventionalized meaning or pragmatic force, and yet seem to be highly probable sequences of words.[1] Similarly, because *like* and *forget* are commonly used as transitive verbs with NP objects, *like the* and *forget the* are collocations (and have been found in aphasic speech where *like to* and *forget to* were the needed forms, Menn and Duffield 2013). However, they are not conventionalized form-function pairings, so they are not constructions.

At any time in the history of a language, collocations may be on their way to becoming constructions, but may not yet have a 'life of their own,' that is, a conventionalized form-function pairing. For example, the English sequence *in the* varies colloquially from [ɪnðə] to [ɪnnə], [ɪnə], [n̥ə], but the choice is a function of style and rate, unlike the obligatory lexicalized translation equivalent German *im* (from *in* + *dem*). So we would say that *in the* in English is a collocation but not a construction, while the German form *im* is a construction.

Recognizing collocations like *in the* or contractions across the NP–VP boundary (***I'm** home! **There's** the book*) as syntagmatic structures allows them a place in speaker's knowledge, whether or not your theory recognizes them as part of grammar. Further, representing diachronically shifting structures as syntagmas (collocations, constructions, or both) does not distort their indeterminate and often conflicting internal hierarchical structures during the course of grammaticization. Traugott (2008) documents, for example, the thousand-year journey of *a lot* from head-modified-by-prepositional-phrase—'NP1 [of NP2]', with the meaning "a share [of land] that falls to a person by chance ('lot') or inheritance", to modifier-followed-by-head—'[NP1 of] NP2', meaning 'a substantial quantity of' NP2. In a model that incorporates gradient usage information, changes in grammar over time may be discrete, but they may also be fuzzy.

[1] While *in the attic* is an instance of the locative prepositional phrase construction and the second phrase contains instances of several constructions, there appears to be no additional conventionalized meaning or use added by the particular lexical items involved.

14.3.4. *Competition and processing complexity revisited*

Psycholinguistic processing was classically taken to deal with two types of tasks: lexical retrieval and ordering the retrieved items in appropriate structures (for production) or parsing the structure the lexical items occur in (for comprehension). But as we have seen, recent research has shown that processing also involves a third major task, which interacts with the first two in both comprehension and production, namely the resolution of competition among alternative lexical items and structures. Competition is thought to be resolved through the physiological process called *spreading activation*. The brain works by neurons sending activating (and sometimes inhibitory) signals to other neurons that they are connected to. Those neurons, in turn, fire if the sum of the signals they receive pushes their own activation level above a threshold, thus spreading the activation further. In language production, the most highly-activated structure or lexical or grammatical morpheme (at the right point in processing) 'wins' the competition and is produced; comprehension works in the same way.

Many of the effects of paradigm structure and syntagmatic context can be understood as influencing processing complexity through their impact on resolving competition among lexical items, affixes, meanings, and possible structural analyses. High sequential probability of morphemes increases the activation from one item in the sequence to another, and pushes a resolution of competition between possible 'next' items in favor of the most likely one. If the most likely next element is in fact a good choice, this 'push' eases production. But sometimes, the most likely form is ungrammatical in the context, has the wrong meaning for the intended message, or is pragmatically inappropriate. In such a case, the speaker must inhibit her impulse to use that form. The effort involved in inhibiting a strongly aroused form makes production more difficult; we are all familiar with the effort required to avoid clichés and excessive professional jargon in our speaking and writing.

14.3.5 *What is a formalism to do?*

We are not the first to suggest that formal descriptions of language need to take into account behaviors observed in psycholinguistic research. From the perspective of psycholinguistics and cognition, a competence–performance distinction "recogniz[es] that [linguistic research] eventually must be placed in a broader psychological context" (Jackendoff 2002), which we take to mean that if linguists are going to assume any sort of psychological reality for a competence grammar, it must eventually be examined using empirical methods. Furthermore, competence grammars are shaped by the experience of the language user (Hawkins 2004). As such, any claims about the link between competence and performance must be backed up by independently motivated, well-defined, empirical research (see also Jaeger and Tily 2011, for a review of psycholinguistic studies).

From the perspective of formal theoretical descriptions, there is already work examining what features a grammar must possess in order to be compatible with measures of performance. For example, Sag and Wasow (2011) argue that performance-compatible grammars must be surface-oriented, constraint-based, and lexicalist for many of the same reasons that we argue for a linguistic complexity metric anchored in the human mind: descriptions of grammar that are responsive to human performance as well as to linguistic structure must take into account competition among syntagmatic and paradigmatic structures and the interaction of information at various linguistic levels.

But the absence of consensus about what parts of the user's information about the language properly belongs in a grammar makes it difficult to frame an argument about whether grammar is amodal with respect to comprehension and production. While constraints on production preferences contribute to the language patterns that provide input for the learner (cf. Gennari and MacDonald 2009), the grammar used in production may be smaller than the grammar used in comprehension, in the same way that one's production vocabulary is smaller than one's comprehension vocabulary. Both follow because any learner is exposed to a variety of inputs that they are not required to match in their output (many of us can read Shakespearean English, but few can write it, for example).

Incorporating syntagmatic structures into grammars poses an additional challenge. Syntagmatic structure (with its nesting constructions and overlapping collocations) contains a great deal of redundancy, and it is further redundant with much of the information contained in hierarchical structures and in the lexicon. Allowing hierarchical structure and syntagmatic structure to co-exist in grammar thus breaks with the tenet that has dominated linguistics since Chomsky and Halle (1968), namely, the idea that the mental representations of linguistic forms should be free of redundancy. But as we have said elsewhere (e.g. Menn and Duffield 2013), neural processing is squeezed for time. Speed and accuracy need to be maximized, so the brain's computations must be robust in noise and massively parallel. This in turn demands redundant storage and processing capacity. And large networks indeed have abundant storage space; brains are not like early-generation computers.

14.4 MISCHA: Model Integrating Sequential, Categorial, and Hierarchical Activation

Because syntagmatic structure and (syntagmatic and paradigmatic) probability distributions affect processing and processing complexity, we are developing 'Model Integrating Sequential, Categorial, and Hierarchical Activation' (MISCHA) to help linguists and psycholinguists to think about the representation of hierarchical grammar and sequential usage simultaneously. MISCHA is a psycholinguistic processing model that should be useful in designing experiments, including experiments that would help to test its principles and further specify its architecture (which is at present only a sketch). At present, MISCHA only models speech production, but the neurally-based principles we are relying on, including connection formation and strengthening, activation, inhibition, and competition resolution, are also applicable to modeling comprehension and learning across modalities.

MISCHA is designed to reflect the following conclusions from the data we have reviewed:

a) Language processing makes use of both hierarchical and syntagmatic (including both constructional and collocational) structures. Human language processing cannot be modeled correctly until the ideal of parsimony is abandoned.
b) Greater frequency (other things being equal) and greater co-occurrence frequency, within and between linguistic levels, creates tighter links between structural elements.
c) The more tightly a morpheme or a structure is linked to a simultaneous or following element (at any linguistic level), the greater the predictability of that element.
d) Complexity is reduced as predictability increases, i.e. as competition for the possible next word or next structure is reduced.

We speak in order to convey messages, and we are constantly projecting what might come next (at least in comprehension and plausibly in production, given our propensity for anticipatory speech errors). So what MISCHA does, in essence, is to unify the classic Garrett-Bock-Levelt message-driven model and Elman's (1990) predictive model. The classic model and the Elman-type prediction models both make valid contributions to our understanding of language production, just as both harmony and melody provide valid contributions to our understanding of music. Here is a quick sketch of how it works.

The Garrett-Bock-Levelt model of on-line sentence (main-clause) production has three principal levels or processing stages (Bock and Levelt 1994; see also Bencini 2013). The first is the Message Level, which creates a well-formed intent-to-say something. This feeds information to the next level 'down', the Functional Level, which is essentially a representation of the message's semantic structure. The Functional Level in turn feeds information to the Positional Level, which specifies the hierarchical constituent structure (the surface structure) of the utterance.

As an on-line production model, MISCHA also starts from the speaker's intent to communicate a particular message, but it takes seriously the idea that the brain is a massively parallel processor and allows the many possible ways of encoding a message to compete with one another. It also incorporates the predictive modeling idea that at every point in the course of sentence production, we make use of past experience to predict and send activation to whatever we are likely to need next. This could be a node, a morpheme or a phoneme.

The key to marrying the classical and the predictive models is the notion of spreading activation. In neural net modeling, it does not matter whether the activation arriving at a neuron arises from disparate sources; all that matters is that it adds up to enough to push that neuron above its threshold for firing.

MISCHA assumes that the items at the three levels of the classic production model (Message Level, Functional Level, Positional Level) arouse their counterparts at the next level down by spreading activation. But also, it assumes that every activated node, at every level, sends activation to the nodes or elements that are likely to co-occur with it or to follow it, implementing the syntagmatic structure we have been arguing for.

The (relatively) static aspects of MISCHA's lexical representation include category-membership information about words (so that they can go into the right slots in morphosyntactic structures), and also, for each word, the frames in which it must or may occur, e.g. the argument frames of each verb. This information is contained in the linkages of networks, and because the linkage strengths are built up by experience, a word's overall frequency, the frequency with which a word is used as a member of a particular category, and the frequency with which it is used in each of its frames (if it has them) are automatically part of this lexical information.

Probability modulates the activation of the competing hierarchical structures as well as the competing syntagmatic structures. The amount of activation that an item sends out is roughly proportional to how likely a particular phrasal expansion or sequential structure is. For example, MISCHA can use the high probability that a noun phrase headed by low-content words like *one* will have a relative clause branch (*The one I was looking for*). The 'common language' for the message encoding and the prediction—the way the two kinds of information interact in the course of sentence production—is activation. At all stages, the interaction between hierarchical and sequential information is represented by summing the activation that they each provide.

MISCHA acquires the probabilistic information that it needs to make predictions by experience, paralleling a child's development. Developmentally, MISCHA is a 'greedy' learner—that is, it picks up any correlation that it happens to find within language and between language and usage (including the communicative setting), without a priori restrictions. Every stored element—every morpheme, word, and structural node—becomes linked with the morphemes, words, and structures that are part of it or that follow it Therefore, it learns and can make use of probabilities of usage (as those of Bybee 2007; Bates and MacWhinney 1987; Saffran 2003; Gennari and MacDonald 2009; Cutler 2012). It also develops links across linguistic levels and between linguistic and real-world knowledge (so that pragmatics can be learned). Probabilistic learning allows it to expand its grammar tentatively, a little at a time, as it encounters structures that it cannot yet generate, making it like Culicover's (1999) 'conservative attentive learner'.

MISCHA's structural representations, like those of the classic production model, are very shallow, but they contain enough information about structure to build the surface components that can be attached to various parts of a sentence as modifiers (principally, in English, adjective phrases, prepositional phrases, relative clauses, and reduced relative clauses). We can think of these structural categories as labeled nodes that reside in MISCHA's network, built up there by linkages across instances of comprehension and production. MISCHA's grammar also contains enough information about noun phrases and verb phrases to specify where modifiers can be attached. This information comes from experience in hearing and/or reading and understanding those structures. MISCHA must also build up enough structural information to identify gaps (e.g. where NP objects fail to follow verbs or prepositions that require them). Because this information is built up from usage and encoded in networks, it automatically reflects the frequency with which a particular structure is used in a particular way (e.g. prepositional phrases as noun post-modifiers or as verb post-modifiers).

MISCHA reflects the way that our brains are constantly making predictions and learning from their successes and failures. Predictions are made by spreading activation. Errors are corrected by reducing the strengths of the connections that led to those errors (or by strengthening inhibitory connections), while correct predictions are reinforced by strengthening the links that created them.

Note that MISCHA's greedy connection-forming habits are not just for mastering pragmatics; they also enable MISCHA to use information that crosses linguistic levels if that information is sufficiently informative and reliable. Not only should MISCHA be able to learn idioms and infer their internal structures, it should be able to build constructions containing specified grammatical morphemes (such as *the* ADJ.X-*er the* ADJ.Y-*er*, Verb X*'s way* (*from* A) *to* B) and to link semantics with a hunk of a hierarchical structure to make general constructions like caused-motion and its metaphorical extensions (*She wired him the money*); cf. Bencini (2013). And nothing keeps MISCHA from learning the occasional correlation between phonological and morphological properties like the English restriction of initial /ð/ to a few subclasses of functors.

14.5 Conclusion: Reviewing the logic

Describing user-independent 'E-language' complexity is essential as a source of hypotheses about complexity-for-users. But creating a psychologically valid measure of complexity of a language as a whole may be impossible, because complexity for the language user is processing complexity, which is dependent on the user's state of knowledge (including the momentary accessibility of that knowledge), and on whether the process is part of comprehension, production, metalinguistic judgment, or some combination of these. Therefore, the complexity of a particular utterance cannot be determined by speaker-independent grammars of any kind. Rather, it requires a processing model that can be adjusted for the speaker's or hearer's knowledge of the language and the world, for whether production or comprehension is under consideration, and also for whether the signal is stable and editable (writing) or fleeting (speech, signing). Psycholinguistics and neurolinguistics can provide a gold standard for testing proposals about specific contributors to complexity-for-users, but not for measuring the complexity of a language.

Experimental and observational data from psycholinguistics and neurolinguistics show that 'usage', namely, syntagmatic information about frequency and co-occurrence frequency, across disparate linguistic levels and major syntactic boundaries, affects language processing complexity. Other things being equal, what is more predictable is simpler. For example, the less paradigmatic competition for a place in the syntagmatic structure, the simpler it is for the user to select one of the competitors. And going along with a statistical bias towards one of the competitors is simpler than going against it.

MISCHA is a conceptual sketch of the kind of processing model that should be able to take these factors into account, and also to support the gradual grammaticization of hierarchy-violating syntagmatic sequences. It is far from a computational or neurological model, but it is a move in those directions compared to its respected parent, the classic Garrett-Bock-Levelt production model.

Acknowledgments

Many thanks for provocation and discussion to Ben Bergen, Giulia Bencini, Joan Bybee, Wally Chafe, Andy Cowell, Peter Culicover, John Dubois, Adele Goldberg, Patrick Juola, Laura Michaelis-Cummings, Marianne Mithun, Bhuvana Narasimhan, Fritz Newmeyer, and the lively audiences at the 2012 University of Washington Conference on Formal Linguistics and the Measurement of Grammatical Complexity and at UC Santa Barbara. We are especially grateful to Fritz and to Laurel Preston for their intensive editorial feedback.

References

Aboh, Enoch O., and Smith, Norval (2009). 'Simplicity, Simplification, Complexity and Complexification: Where Have the Interfaces Gone?', in Enoch O. Aboh and Norval Smith (eds.), *Complex Processes in New Languages*. Amsterdam: John Benjamins, 1–25.

Abusch, Dorit (1985). *On Verbs and Time*. Ph.D. dissertation, University of Massachusetts, Amherst.

Adger, David (2013). *A Syntax of Substance*. Cambridge, MA: MIT Press.

Aikhenvald, Alexandra Y., and Dixon, R. M. W. (1998). 'Dependencies between Grammatical Systems', *Language* 74: 56–80.

Aitchison, Jean (2000). *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.

Akmajian, Andrew, Demers, Richard A., and Harnish, Robert M. (1979). *Linguistics: An Introduction to Language and Communication*. Cambridge, MA: MIT Press.

Aldridge, Edith (2004). *Ergativity and Word Order in Austronesian Languages*. Ph. D. dissertation, Cornell University.

Alexiadou, Artemis (2001). *Functional Structure in Nominals: Nominalization and Ergativity*. Amsterdam: John Benjamins.

Alexiadou, Artemis, and Stavrou, Melita (1998). 'On Derived Nominals in Greek', in B. Joseph, G. Horrocks, and I. Philippaki-Warbarton (eds.), *Themes In Greek Linguistics II*. Amsterdam: John Benjamins, 101–129.

Allen, Cynthia L. (1980). 'Movement and Deletion in Old English', *Linguistic Inquiry* 11: 261–323.

Allwood, Jens (1982). 'The Complex NP Constraint in Swedish', in E. Engdahl and E. Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages*. Stockholm: Almqvist and Wiksell, 15–32.

Anderson, Stephen (1976). 'On the Notion of Subject in Ergative Languages', in C. N. Li (ed.) *Subject and Topic*. New York: Academic Press, 1–24.

Anderson, Stephen (2004). *Doctor Dolittle's Delusion: Animals and the Uniqueness of Human Language*. New Haven: Yale University Press.

Ansaldo, Umberto, and Matthews, Stephen J. (2007). 'Deconstructing Creole, The Rationale', in U. Ansaldo and S.J. Matthews (eds.), *Deconstructing Creole: New Horizons in Language Creation*. Amsterdam: John Benjamins, 1–18.

Ansaldo, Umberto, and Nordhoff, Sebastian (2009). 'Complexity and the Age of Languages', in Enoch O. Aboh and Norval Smith (eds.), *Complex Processes in New Languages*. Amsterdam: John Benjamins, 345–363.

Arce-Arenales, Manuel, Axelrod, Melissa, and Fox, Barbara A. (1994). 'Active Voice and Middle Diathesis: A Cross-Linguistic Perspective', in B. Fox and P. J. Hopper (eds.), *Voice: Form and Function*. *Typological Studies in Language* 27. Amsterdam: John Benjamins, 1–21.

Arends, Jacques (2001). 'Simple Grammars, Complex Languages', *Linguistic Typology* 5: 180–182.

Ariel, Mira (1990). *Accessing Noun-Phrase Antecedents*. London: Routledge.

Armbruster, Carl Hubert (1960). *Dongolese Nubian: A Grammar*. Cambridge: Cambridge University Press.

Arnaiz, Alfredo, and Camacho, José (1999). 'A Topic Auxiliary in Spanish', in J. Gutiérrez-Rexach, and F. Martínez-Gil (eds.), *Advances in Hispanic Linguistics: Papers from the 2nd Hispanic Linguistics Symposium*, volume 2. Somerville, MA: Cascadilla, 317–331.

Aronoff, Mark, Meir, Irit, Padden, Carol, and Sandler, Wendy (2008). 'The Roots of Linguistic Organization in a New Language', in Derek Bickerton and Michael Arbib (eds.), *Holophrasis, Compositionality and Protolanguage* (Special issue of *Interaction Studies*), 133–149.

Aronoff, Mark, Meir, Irit and Sandler, Wendy (2000). 'Universal and Particular Aspects of Sign Language Morphology', in K.K. Grohmann and C. Struijke (eds.), *University of Maryland Working Papers in Linguistics* 10: 1–34.

Aronoff, Mark, Meir, Irit and Sandler, Wendy (2005). 'The Paradox of Sign Language Morphology', *Language* 81: 301–334.

Assmann, Anke, Georgi, Doreen, Heck, Fabian, Müller, Gereon, and Weisser, Phillip (2012). 'Ergatives Move Too Early: On an Instance of Opacity in Syntax'. Unpublished MS, http://ling.auf.net/lingbuzz/001646 (Accessed 5th March 2014).

Atkinson, Quentin D. (2011). 'Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa', *Science* 332: 346–359.

Austin, Peter (1981). *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.

Baayen, R. Harald (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bach, Emmon, Brown, Colin, and Marslen-Wilson, William (1986). 'Crossed and Nested Dependencies in German and Dutch: A Psycholinguistic Study', *Language and Cognitive Processes* 1: 249–262.

Bach, Emmon, Jelinek, Eloise, Kratzer, Angelika and Partee, Barbara (eds.) (1995). *Quantification in Natural Languages*. Dordrecht: Kluwer.

Badan, Linda, and Del Gobbo, Francesca (2011). 'On the Syntax of Topic and Focus in Chinese', in P. Benincà and N. Munaro (eds.), *Mapping the Left Periphery. The Cartography of Syntactic Structure* (volume 5). Oxford: Oxford University Press, 63–90.

Bader, Markus, and Schmid, Tanja (2009a). 'Cat Meets Go: 2-Verb Clusters in German', in Jeroen van Craenenbroeck (ed.), *Alternatives to Cartography*. Berlin: Mouton de Gruyter, 203–245.

Bader, Markus, and Schmid, Tanja (2009b). 'Minimality in Verb-Cluster Formation', *Lingua* 119: 1458–1481.

Bader, Markus, and Schmid, Tanja (2009c). 'Verb Clusters in Colloquial German', *The Journal of Comparative Germanic Linguistics* 12: 175–228.

Baerman, Matthew, Brown, Dunstan, and Corbett, Greville G. (eds.) (2014). *Understanding and Measuring Morphological Complexity*. Oxford: Oxford University Press.

Baird, Louise (2008). *A Grammar of Klon: A Non-Austronesian Language of Alor, Indonesia*. Canberra: Australian National University.

Baker, Mark (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Baker, Mark (1996). *The Polysynthesis Parameter*. New York: Oxford University Press.

Baker, Mark (2001). *The Atoms of Language: The Mind's Hidden Rules of Grammar*. New York: Basic Books.

Bakker, Peter (1997). *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French Language of the Canadian Metis*. New York: Oxford University Press.

Bakker, Peter (2012). 'Plains Indian Sign Languge: Structure of a Signed Pidgin', Paper presented at the Ninth Creolistics Workshop, Aarhus, 11 April 2012.

Bakker, Peter, Daval-Markussen, Aymeric, Parkvall, Mikael and Plag, Ingo (2011). 'Creoles Are Typologically Distinct from Non-Creoles', *Journal of Pidgin and Creole Languages* 26: 5–42.

Bane, Max (2008). 'Quantifying and Measuring Morphological Complexity', in Charles B. Chang and Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, 69–76.

Barwise, Jon, and Cooper, Robin (1981). 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* 4: 159–219.

Bastiaanse, Roelien, Jonkers, R., Ruigendijk E., and van Zonneveld, R. (2003). 'Gender and Case in Agrammatic Production', *Cortex* 39: 405–417.

Bates, Elizabeth, and MacWhinney, Brian (1987). 'Competition, Variation and Language Learning', in B. MacWhinney (ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum, 157–194.

Bates, Elizabeth, Wulfeck, Beverley, and MacWhinney, Brian (1991). 'Crosslinguistic Research in Aphasia: An Overview', *Brain and Language* 41: 123–148.

Battistella, Edward L. (1996). *The Logic of Markedness*, New York: Oxford University Press.

Beebe, Ralph D. (1973). 'The Frequency Distribution of English Syntagms', in A. Zampolli, and N. Calzolari (eds.), *COLING '73: Proceedings of the 5th Conference on Computational Linguistics*, volume 2. Stroudsburg, PA: Association for Computational Linguistics, 9–22.

Belletti, Adriana, Friedmann, Naama, Brunato, Dominique and Rizzi, Luigi (2012). 'Does Gender Make a Difference? Comparing the Effect of Gender on Children's Comprehension of Relative Clauses in Hebrew and Italian', *Lingua*, 122: 1053–1069.

Belletti, Adriana, and Rizzi, Luigi (2013). 'Intervention in Grammar and Processing', in Ivano Caponigro and Carlo Cecchetto (eds.), *From Grammar to Meaning: The Spontaneous Logicality of Language*. Cambridge: Cambridge University Press, 293–311.

Ben-Shachar, Michal, Palti, Dafna, and Grodzinsky, Yosef (2004). 'Neural Correlates of Syntactic Movement: Converging Evidence from two fMRI Experiments', *Neuroimage* 21: 1320–1336.

Bencini, Giulia M. L. (2013). 'Psycholinguistics', in Thomas Hoffmann and Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.

Benedetto, Dario, Caglioti, Emanuele, and Loreto, Vittorio (2002). 'Language Trees and Zipping', *Physical Review Letters* 88: 48702.

Bergman, Elizabeth M. (2002). *Spoken Sudanese Arabic, Grammar, Dialogues, and Glossary*. Springfield VA: Dunwoody Press.

Berndt, Rita Sloan (2001). 'Sentence Production', in Brenda Rapp (ed.) *Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind.* New York: Psychology Press.

Bernstein, Judy (1991). 'DPs in French and Walloon: Evidence for Parametric Variation in Nominal Head Movement', *Probus* 3: 101–126.

Berwick, Robert (1985). *The Acquisition of Syntactic Knowledge.* Cambridge, MA: MIT Press.

Berwick, Robert, and Chomsky, Noam (2011). 'The Biolinguistic Program: The Current State of its Development', in A. M. Di Sciullo and C. Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford: Oxford University Press, 19–41.

Bever, Thomas (1970). 'The Cognitive Basis for Linguistic Structures', in John R. Hayes (ed.), *Cognition and the Development of Language*. New York: John Wiley and Sons.

Bianchi, Valentina (2002). 'Headed Relative Clauses in Generative Syntax. Part I', *Glot International* 6: 197–204.

Bianchi, Valentina (2009). 'A Note on Backward Anaphora', *Rivista Di Grammatica Generativa* 34: 3–34.

Biberauer, Theresa (n.d.) 'Disharmonic Word Order, Quirky Morphology and the Afrikaans Verb Cluster', Available online at http://www.hum.au.dk/engelsk/engsv/nyvad-abstracts/biberauer-ho.pdf (Accessed 5th March 2014).

Biberauer, Theresa (2011). 'In Defence of Lexico-Centric Parametric Variation: Two 3rd Factor-Constrained Case Studies'. Paper presented at the Workshop on Formal Grammar and Syntactic Variation: Rethinking Parameters (Madrid).

Biberauer, Theresa (2013). 'Offers That Cannot be Refused: Mafioso Parameters and the Construction of Minimalist Parameters'. Paper presented at the 39th Incontro Di Grammatica Generativa (Modena and Reggio Emilia).

Biberauer, Theresa, Holmberg, Anders, and Roberts, Ian (2014). 'A Syntactic Universal and Its Consequences', *Linguistic Inquiry* 45(2): 169–225.

Biberauer, Theresa, Holmberg, Anders, Roberts, Ian and Sheehan, Michelle (2010). 'Reconciling Formalism and Functionalism: A Minimalist Perspective'. Paper presented at the LAGB Annual Meeting (Leeds).

Biberauer, Theresa, and Richards, Marc. (2006). 'True Optionality: When the Grammar Doesn't Mind', in C. Boeckx (ed.), *Minimalist Essays*. Amsterdam: John Benjamins, 35–67.

Biberauer, Theresa, and Roberts, Ian (2012a). 'Towards a Parameter Hierarchy for Auxiliaries: Diachronic Considerations', in J. Chancharu, X. Hu and M. Mitrović (eds.), *Cambridge Occasional Papers in Linguistics* 6: 209–36.

Biberauer, Theresa, and Roberts, Ian (2012b). 'On the Significance of What Doesn't Happen'. Paper presented at the 14th Diachronic Generative Syntax (DIGS 14) Conference (Lisbon).

Biberauer, Theresa, and Roberts, Ian (2013). 'Clausal Hierarchies'. Unpublished MS, University of Cambridge.

Biberauer, Theresa, Roberts, Ian and Sheehan, Michelle (2014). 'No-Choice Parameters and the Limits of Syntactic Variation'. To appear in N. Danton, D. Kostadinovska and R. Santana-Labarge (eds.), *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL 31).* Somerville, MA: Cascadilla.

Bickel, Balthasar (2008). *A General Method for the Statistical Evaluation of Typological Distributions*. MS, University of Leipzig. http://www.spw.uzh.ch/bickel-files/papers/testing_universals_bickel2008.pdf (Accessed 5th March 2014).

Bickel, Balthasar (in press). 'Linguistic Diversity and Universals', in Nick Enfield, Paul Kockelman, and Jack Sidnell (eds.), *The Cambridge Handbook of Linguistic Anthropology*. Cambridge: Cambridge University Press.

Bickel, Balthasar, and Nichols, Johanna (2007). 'Inflectional Morphology', in T. Shopen (ed.), *Language Typology and Syntactic Description,* vol. 3: *Grammatical Categories and the Lexicon* (2nd edn.). Cambridge: Cambridge University Press, 169–240.

Bickel, Balthasar, and Nichols, Johanna (2011). 'Inflectional Synthesis of the Verb', in Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, ch. 22. http://wals.info/chapter/22 (Accessed 5th March 2014).

Bickerton, Derek (1977). 'Pidginization and Creolization: Language Acquisition and Language Universals', in A. Valdman (ed.), *Pidgin and Creole Linguistics*. Bloomington: Indiana University Press, 49–69.

Bickerton, Derek (1981). *Roots of Language*. Ann Arbor: Karoma.

Bickerton, Derek (1984). 'The Language Bioprogram Hypothesis', *Behavioral and Brain Sciences* 7: 173–221.

Bickerton, Derek (1990). *Language and Species*. Chicago: University of Chicago Press.

Bickerton, Derek (1995). *Language and Human Behavior*. Seattle: University of Washington Press.

Bickerton, Derek (1998). 'Catastrophic Evolution: The Case for a Single Step from Protolanguage to Full Human Language', in J.R. Hurford, M. Studdert-Kennedy, and C. Knight (eds.), *Approaches to The Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press, 341–358.

Bickerton, Derek (2008). *Bastard Tongues*. New York: Hill and Wang.

Biedermann, Britta, Lorenz, Antje, Beyersmann, Elisabeth, and Nickels, Lyndsey (2012). 'The Influence of Plural Dominance in Aphasic Word Production', *Aphasiology* 26: 985–1004.

Bies, Ann E. (1996). *Syntax and Discourse Factors in Early New High German: Evidence for Verb-Final Word Order,* MA thesis, University of Pennsylvania.

Bisang, Walter (2009). 'On the Evolution of Complexity: Sometimes Less is More in East and Mainland Southeast Asia', in Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 34–49.

Blake, Barry J. (2001). *Case* (2nd edn.). Cambridge: Cambridge University Press.

Bloom, Lois (1970). *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.

Boas, Franz (1963 [1911]). *Introduction to the Handbook of American Indian Languages*. Washington: Georgetown University Press.

Bobaljik, Jonathan D. (1995). 'In Terms of Merge: Copy and Head Movement', *MIT Working Papers in Linguistics* 27: 41–64.

Bock, Kathryn, and Cutting, J. Cooper (1992). 'Regulating Mental Energy: Performance Units in Language Production', *Journal of Memory and Language* 31: 99–127.

Bock, Kathryn, Irwin, David E., and Davidson, Douglas J. (2004). 'Putting First Things First', in John Michael Henderson and Fernanda Ferreira (eds). *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*. New York: Psychology Press.

Bock, Kathryn, and Levelt, Willem J. M. (1994). 'Language Production: Grammatical Encoding', in M. A. Gernsbacher (ed.), *Handbook of Psycholinguistics*. San Diego, CA: Academic Press, 945–984.

Boeckx, Cedric (2006). *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.

Bok-Bennema, Reineke (1991). *Case and Agreement in Inuit*. Dordrecht: Foris.

Borer, Hagit (1994). 'The Projection of Arguments', in E. Benedicto and J. Runner (eds.), *University of Massachusetts Occasional Papers in Linguistics* 17: 19–47.

Bornkessel-Schlesewsky, Ina, and Schlesewsky, Matthias (2009). 'Minimality as Vacuous Distinctness: Evidence from Cross-Linguistic Sentence Comprehension', *Lingua* 119: 1541–1559.

Borsley, Robert D., Tallerman, Maggie, and Willis, David (2007). *The Syntax of Welsh*. Cambridge: Cambridge University Press.

Bošković, Željko (1995). *Principles of Economy in Nonfinite Complementation*. Ph.D. dissertation, University of Connecticut. Published in 1997 as *The Syntax of Nonfinite Complementation: An Economy Approach* by MIT Press, Cambridge, MA.

Bošković, Željko (2008). 'What Will You Have, DP or NP?' in M. Walkow and E. Elfner (eds.), *Proceedings of NELS* 37: 101–114.

Bott, Lewis, and Noveck, Ira A. (2004). 'Some Utterances are Underinformative: The Onset and Time Course of Scalar Inferences, *Journal of Memory and Language* 51: 437–457.

Bottari, Piero (1992). 'Romance Passive Nominals', *Geneva Generative Papers*, 1: 66–80.

Bouma, Gosse, and van Noord, Gertjan (1998). 'Word Order Constraints on Germanic Verb Clusters', in Erhard Hinrichs, Tsuneko Nakazawa, and Andreas Kathol (eds.), *Complex Predicates in Nonderivational Syntax.* Orlando: Academic Press.

Bradshaw, Joel (2009). 'Relative-Clause Bracketing in Oceanic Languages around the Huon Gulf of New Guinea', in A. Adelaar and A. Pawley (eds.), *Austronesian Historical Linguistics and Culture History. A Festschrift for Robert Blust*. Canberra: Pacific Linguistics, 143–162.

Braine, Martin (1963). 'The Ontogeny of English Phrase Structure: The First Phase', *Language* 39: 11–14.

Branigan, Phil (2011). 'In Defense of Macroparameters: Algonquian Evidence'. Paper presented at the Workshop on Formal Grammar and Syntactic Variation: Rethinking Parameters (Madrid).

Branigan, Phil (2012). 'Macroparameter Learnability: An Algonquian Case Study'. Unpublished MS, Memorial University of Newfoundland.

Breedin, Sarah D., Saffran Eleanor M., and Schwartz Myrna F. (1998). 'Semantic Factors in Verb Retrieval: An Effect of Complexity', *Brain and Language* 63: 1–31.

Breheny, Richard, Katsos, Napoleon, and Williams, John (2006). 'Are Generalized Scalar Implicatures Generated By Default? An On-Line Investigation into the Role of Context in Generating Pragmatic Inferences', *Cognition* 100: 434–463.

British National Corpus, version 3 (BNC XML Edition) (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Available online at: http://www.natcorp.ox.ac.uk/ (Accessed 5th March 2014).

Broadbent, Sylvia M. (1964). *The Southern Sierra Miwok Language*. Berkeley, CA: University of California Press.

Broadwell, George Aaron (2006). *Choctaw Reference Grammar*. Lincoln, NB: University of Nebraska Press.

Brody, Michael (2002). 'On the Status of Representations and Derivations', in S. D. Epstein and T. D. Seely (eds.), *Derivation and Explanation in the Minimalist Program*. Malden, MA: Blackwell, 19–41.

Brown, Keith, and Miller, Jim (1996). *Concise Encyclopedia of Syntactic Theories*. Oxford: Elsevier Science.

Brown, Lea (2001). *A Grammar of Nias Selatan*. Ph.D. dissertation, University of Sydney.

Brown, Roger (1973). *A First Language.* Cambridge, MA: Harvard University Press.

Burling, Robbins (2005). *The Talking Ape: How Language Evolved*. Oxford: Oxford University Press.

Burzio, Luigi (1981). *Intransitive Verbs and Italian Auxiliaries*. Ph.D. dissertation, MIT.

Bybee, Joan L. (2007). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

Bybee, Joan L., and Thompson, Sandra A. (1997). 'Three Frequency Effects in Syntax', in *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 65–85.

Campbell, George (1991). *Compendium of the World's Languages*. London: Routledge.

Caplan, David (2001). 'Functional Neuroimaging Studies of Syntactic Processing', *Journal of Psycholinguistic Research* 30: 297–320.

Caplan, David, Alpert, Nathaniel, and Waters, Gloria S. (1998). 'Effects of Syntactic Structure and Propositional Number on Patterns of Regional Cerebral Blood Flow', *Journal of Cognitive Neuroscience* 10: 541–552.

Cappa, Stefano (2012). 'Imaging Semantics and Syntax', *Neuroimage* 61: 427–431.

Caramazza, Alfonso, and Zurif, Edgar (1976). 'Dissociation of Algorithmic and Heuristic Processes in Language Comprehension: Evidence from Aphasia', *Brain and Language* 3: 572–582.

Carden, Guy, and Pesetsky, David, (1977). 'Double-Verb Construction, Markedness, and a Fake Coordination', in W. A. Beach, S. E. Fox, and S. Philosoph (eds.), *Papers from the Thirteenth Regional Meeting of the Chicago Linguistics Society*. Chicago, IL: Chicago Linguistic Society, 82–92.

Cardoso, Hugo C. (2009). *The Indo-Portuguese Language of Diu*. Utrecht: LOT.

Carlson, Robert (1994). *A Grammar of Supyire*. Berlin: Walter de Gruyter.

Carroll, John B., and White, Margaret N. (1973). 'Word Frequency and Age of Acquisition as Determiners of Picture-Naming Latency', *Quarterly Journal of Experimental Psychology* 25: 85–95.

Carroll, Sean B. (2005). *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom*. New York: W. W. Norton and Company.

Casad, Eugene H. (1984). 'Cora', in Ronald W. Langacker (ed.), *Studies in Uto-Aztecan Grammar 4: Southern Uto-Aztecan Grammatical Sketches*. Dallas and Arlington, TX: Summer Institute of Linguistics and the University of Texas, 153–459.

Casielles, Eugenia, and Progovac, Ljiljana (2010). 'On Protolinguistic "Fossils": Subject-Verb vs. Verb-Subject Structures', in A. D. M. Smith, M. Schouwstra, B. De Boer, and K. Smith, (eds.), *The Evolution of Language: Proceedings of The 8th International Conference (EVOLANG8)*. New Jersey: World Scientific, 66–73.

Casielles, Eugenia, and Progovac, Ljiljana (2012). 'Protosyntax: A Thetic (Unaccusative) Stage', *Theoria Et Historia Scientiarum: An International Journal for Interdisciplinary Studies* IX: 29–48.

Chaitin, Gregory J. (1987). *Algorithmic Information Theory*. Cambridge: Cambridge University Press.

Chater, Nick, and Vitány, Paul (2007). 'Ideal Learning of Natural Language: Positive Results about Learning from Positive Evidence', *Journal of Mathematical Psychology* 51: 135–63.

Cheng, Lisa, and Rint, Sybesma (2003). 'Forked Modality', *Linguistics in the Netherlands,* 13–23.

Chesi, Cristiano (2007). 'An Introduction to Phase-based Minimalist Grammars: Why 'Move' is Top-down from Left-to-right', *CISCL Working Papers on Language and Cognition* 1: 38–75.

Chesi, Cristiano (2014). 'What Memory Retrieval Tells Us About Intervention Effects', in Abstract Proceedings of IGG40 (Italian Generative Grammar Meeting), Trento, 13–15 February 2014, 50–51. http://events.unitn.it/sites/events.unitn.it/files/download/igg40/igg40-book-of-abstracts.pdf.

Chierchia, Gennaro, Fox, Danny, and Spector, Benjamin (2012). 'The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics', in Paul Portner, Claudia Maienborn, and Klaus von Heusinger (eds.), *Handbook of Semantics*, vol. 3. Berlin: Mouton de Gruyter.

Childs, G. Tucker (1995). *A Grammar of Kisi: A Southern Atlantic Language*. Berlin: Mouton de Gruyter.

Chomsky, Noam (1956). 'Three models for the description of language', *IRE Transactions on Information Theory* 2: 113–124.

Chomsky, Noam (1957). *Syntactic Structures.* The Hague: Mouton.

Chomsky, Noam (1959). 'On Certain Formal Properties of Grammars', *Information and Control* 2: 137–167.

Chomsky, Noam (1963). 'Formal Properties of Grammars', in R. D. Luce, R. R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*: vol. 2. New York: Wiley, 323–418.

Chomsky, Noam (1964). *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam (1973). 'Conditions on Transformations', in S. Anderson and P. Kiparsky (eds.). *A Festschrift for Morris Halle*. New York: Holt, Reinhart and Winston, 232–86.

Chomsky, Noam (1975 [1955]). *The Logical Structure of Linguistic Theory*. New York and London: Plenum Press.

Chomsky, Noam (1976). 'Conditions on Rules of Grammar', *Linguistic Analysis* 2: 303–351.

Chomsky, Noam (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Chomsky, Noam (1991). 'Some Notes on Economy of Derivation and Representation', in R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*. Cambridge, MA: MIT Press, 417–454.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000). 'Minimalist Inquiries: The Framework', in R. Martin, D. Michaels and J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press, 89–156.

Chomsky, Noam (2001). 'Derivation by Phase', in M. Kenstowicz (ed.). *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1–53.

Chomsky, Noam (2004a). *The Generative Enterprise Revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui, and Mihoko Zushi*. Berlin and New York: Mouton de Gruyter.

Chomsky, Noam (2004b). 'Beyond Explanatory Adequacy', in Adriana Belletti (ed.), *Structures and Beyond*. Oxford: Oxford University Press, 104–131.

Chomsky, Noam (2005). 'Three Factors in Language Design', *Linguistic Inquiry* 36: 1–22.

Chomsky, Noam (2007). 'Approaching UG From Below', in U. Sauerland and H.-M. Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View From Syntax-Semantics*. Berlin: Mouton De Gruyter, 1–29.

Chomsky, Noam (2013). 'Problems of Projection', *Lingua* 130: 33–49.

Chomsky, Noam, and Halle, Morris (1968). *The Sound Pattern of English*. New York: Harper and Row.

Chomsky, Noam, and Miller, George A. (1963). 'Introduction to The Formal Analysis of Natural Languages', in R.D. Luce, R.R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, vol. 2. New York: Wiley, 269–321.

Christiansen, Morten H., and Chater, Nick (1999). 'Toward a Connectionist Model of Recursion in Human Linguistic Performance', *Cognitive Science* 23: 157–205.

Christiansen, Morten H., and MacDonald, Maryellen C. (2009). 'A Usage-based Approach to Recursion in Sentence Processing', *Language Learning* 59: 126–161.

Chung, Sandra, and Ladusaw, William A. (2004). *Restriction and Saturation*. Cambridge, MA: MIT Press.

Cinque, Guglielmo (1996). 'The "Antisymmetric" Program: Theoretical and Typological Implications', *Journal of Linguistics* 32: 447–65.

Citko, Barbara (2005). 'On the Nature of Merge: External Merge, Internal Merge, and Parallel Merge', *Linguistic Inquiry* 36: 475–497.

Clements, G. N. (2009). 'The Role of Features in Phonological Inventories', in Eric Raimy and Charles E. Cairns (eds.), *Contemporary Views on Architecture and Representations in Phonology*. Cambridge, MA: MIT Press, 19–68.

Code, Chris (2005). 'First In, Last Out? The Evolution of Aphasic Lexical Speech Automatisms to Agrammatism and the Evolution of Human Communication', *Interaction Studies* 6: 311–334.

Cohen, Kevin Bretonnel (2000). *Aspects of the Grammar of Kukú*. Munich: Lincom Europa.

Colapinto, John (2007). 'Review of D. Everett, *Don't Sleep, There are Snakes*', *New Yorker*, 16th April 2007.

Cole, Peter (1985). *Imbabura Quechua*. London: Croom Helm.

Collins, Chris (1997). *Local Economy*. Cambridge, MA: MIT Press.

Comrie, Bernard (1973). 'Clause Structure and Movement Constraints in Russian', in C. Corum, T.C. Smith-Stark and A. Weiser (eds.), *You Take the High Node and I'll Take The Low Node*. Chicago: Chicago Linguistic Society, 291–304.

Comrie, Bernard (1978). 'Ergativity', in W. P. Lehmann (ed.), *Syntactic Typology: Studies in the Phenomenology of Language*. Austin: University of Texas Press, 329–394.

Comrie, Bernard (1989). *Language Universals and Linguistic Typology: Syntax and Morphology* (2nd edn.). Oxford: Blackwell.

Comrie, Bernard (2011). 'Alignment of Case Marking', in Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, ch. 98. http://wals.info/chapter/98 (Accessed 5th March 2014).

Conrad, Robert J., and Wogiga, Kepas (1991). *An Outline of Bukiyip Grammar*. Canberra: Australian National University.

Coon, Jessica, Mateo, Pedro and Preminger, Omer (2011). 'The Role of Case in A-Bar Extraction Asymmetries: Evidence from Mayan'. Unpublished MS, McGill, Harvard and Syracuse Universities. http://people.linguistics.mcgill.ca/~Jessica/Papers_And_Handouts_Files/Coonmateopreminger-AF.pdf (Accessed 5th March 2014).

Cornips, Leonie (2009). 'Empirical Syntax: Idiolectal Variability in Two-and Three-Verb Clusters in Regional Standard Dutch and Dutch Dialects', in Andreas Dufter, Jürg Fleische, and Guido Seiler (eds.), *Describing and Modeling Variation in Grammar*. Berlin, New York: Mouton De Gruyter.

Coseriu, Eugenio (1966). '"Tomo y me voy": ein Problem vergleichender europäischer Syntax', *Vox Romanica* 25: 13–55.

Cover, Rebecca (2011). 'Modal Aspects of Badiaranke Aspect', *Lingua* 121: 1315–1343.

Cresswell, Cassandre, and Snyder, Kieran (2000). 'Passive and Passive-Like Constructions in Hmong', in Roger Billerey and Brook Danielle Lillehaugen (eds.), *Proceedings of the 19th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, 71–82.

Croft, William (2007). 'Construction Grammar', in D. Geeraerts and H. Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, 463–508.

Crothers, John H., Lorentz, James P., Sherman, Donald A., and Vihman, Marilyn M. (1979). *Handbook of Phonological Data from a Sample of the World's Languages: A Report of the Stanford Phonology Archive*. Stanford: Stanford Phonological Archive.

Crystal, David (1987). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

Culicover, Peter W. (1999). *Syntactic Nuts: Hard Cases, Syntactic Theory, and Language Acquisition*. New York: Oxford University Press.

Culicover, Peter W. (2008). 'The Rise and Fall of Constructions and the History of English *Do*-Support', *Journal of Germanic Linguistics* 20: 1–52.

Culicover, Peter W. (2011). 'A Reconsideration of English Relative Constructions', *Constructions*, 2: 1–14.

Culicover, Peter W. (2013). *Grammar and Complexity: Language at the Intersection of Competence and Performance*. Oxford: Oxford University Press.

Culicover, Peter W., and Jackendoff, Ray (2005). *Simpler Syntax*. Oxford: Oxford University Press.

Culicover, Peter W., and Jackendoff, Ray (2006). 'The Simpler Syntax Hypothesis', *Trends in Cognitive Sciences* 10: 413–418.

Culicover, Peter W., and Nowak, Andrzej (2003). *Dynamical Grammar*. Oxford: Oxford University Press.

Curry, Haskell B. (1961). 'Some Logical Aspects of Grammatical Structure', in Roman Jakobson (ed.), *Structure of Language and Its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics*. American Mathematical Society.

Cutler, Anne (2012). *Native Listening: Language Experience and the Recognition of Spoken Words.* Cambridge, MA: MIT Press.

Cutting, J. Cooper, and Bock, J. Kathryn (1997). 'That's the Way the Cookie Bounces: Syntactic and Semantic Components of Experimentally Elicited Idiom Blends', *Memory and Cognition* 25: 57–71.

Cysouw, Michael (2007). 'On the probability distribution of typological frequencies', in *Proceedings of the 10th and 11th Biennial Conference on the Mathematics of Language*, 29–35.

Cysouw, Michael, Dediu, Dan, and Moran, Steven (2012). 'Still No Evidence for an Ancient Language Expansion From Africa', *Science* 335: 657.

Dahl, Östen (2004). *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.

Dahl, Östen (2008). 'Grammatical Resources and Linguistic Complexity: Sirionó as a Language Without NP Coordination', in Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 153–64.

Dahl, Östen (2009). 'Testing the Assumption of Complexity Invariance: The Case of Elfdalian and Swedish', in G. Sampson, D. Gil, and P. Trudgill (eds.), *Language Complexity as an Evolving Variable*. New York: Oxford University Press, 50–63.

Dahlstrom, Amy (1991). *Plains Cree Morphosyntax*. New York and London: Garland Publishing Inc.

Danforth, Bryan N., Conway, Lindsay, and Ji, Shuqing (2003). 'Phylogeny of Eusocial Lasioglossum Reveals Multiple Losses of Eusociality within a Primitively Eusocial Clade of Bees (Hymenoptera: Halictidae)', *Systems Biology* 52: 23–36.

Dapretto, Mirella, and Bookheimer, Susan Y. (1999). 'Form and Content: Dissociating Syntax and Semantics in Sentence Comprehension', *Neuron*, 242: 427–432.

Davies, Mark (2008–). *The Corpus of Contemporary American English: 450 million words, 1990-present.* http://corpus.byu.edu/coca/ (Accessed 5th March 2014).

Davis, Cathlin M. (2002). *Case and Pronominal Suffixes in Sierra Miwok*. Ph.D. dissertation, University of Wisconsin-Madison.

Davis, Henry (2010). 'Salish Languages Lack Generalized Quantifiers After All!' Paper presented at SALT 20, University of British Columbia and Simon Fraser University.

Davis, Henry (2012). *A Teaching Grammar of Upper St'át'imcets*. MS, University of British Columbia.

Davis, Henry, Gillon, Carrie, and Matthewson, Lisa (in press). 'How to Investigate Linguistic Diversity: Lessons From the Pacific Northwest', *Language (Perspectives)*.

Davis, Henry, Matthewson, Lisa, and Rullmann, Hotze (2009). 'Out of Control Marking as Circumstantial Modality in St'át'imcets', in Lotte Hogeweg, Helen de Hoop and Andrey Malchukov (eds.), *Cross-Linguistic Semantics of Tense, Aspect and Modality*. Oxford: John Benjamins, 206–244.

Dayley, Jon P. (1985). *Tzutujil Grammar*. Berkeley, CA: University of California Press.

De Belder, Marijke, and Craenenbroeck, Jeroen van (2011). 'How to Merge a Root', MS, CRISSP Brussels. http://www.crissp.be/wp-content/uploads/papermjroots.pdf (Accessed 5th March 2014).

de Quadros, Müller, Ronice and Lillo-Martin, Diane (2010). 'Clause Structure', in D. Brentari (ed.), *Sign Languages*. Cambridge: Cambridge University Press, 225–251.

de Vos, Mark (2005). *The Syntax of Verbal Pseudo-coordination in English and Afrikaans.* Utrecht: LOT.

De Weerdt, Danny, and Vermeerbegen, Myriam (2008). 'Observations on Possessive and Existential Constructions in Flemish Sign Language', in U. Zeshan and P. Perniss (eds.), *Possessive and Existential Constructions in Sign Languages*, 195–211.

Deacon, Terrence (1997). *The Symbolic Species*. New York: Norton.

Deal, Amy Rose (2011). 'Modals without scales'. *Language* 87: 559–585.

DeGraff, Michel (2001). 'On the Origin of Creoles: A Cartesian Critique of Neo-Darwinian Linguistics', *Linguistic Typology* 5: 213–310.

DeGraff, Michel (2003). 'Against Creole Exceptionalism', *Language* 79: 391–410.

DeGraff, Michel (2005). 'Linguists' Most Dangerous Myth: The Fallacy of Creole Exceptionalism', *Language in Society* 34: 533–591.

Demberg, Vera, and Keller, Frank (2008). 'Data from Eye-Tracking Corpora as Evidence for Theories of Syntactic Processing Complexity', *Cognition*, 109: 193–210.

Dennett, Daniel C. (1995). *Darwin's Dangerous Idea*. New York: Simon and Schuster.

Derbyshire, Desmond C. (1979). *Hixkaryana*. Amsterdam: North-Holland.

Derbyshire, Desmond C. (1985). *Hixkaryana and Linguistic Typology*. Dallas, TX: Summer Institute of Linguistics.

Deutscher, Guy (2005). *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest Invention*. New York: Metropolitan Books.

Deutscher, Guy (2009). ''Overall Complexity'—a Wild Goose Chase?', in Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 243–251.

Di Sciullo, Anna Maria, and Boeckx, Cedric (eds.) (2011). *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford: Oxford University Press.

Dick, Frederic, Bates, Elizabeth, Wulfeck, Beverly, Utman, Jennifer A., Dronkers, Nina, and Gernsbacher, Morton Ann (2001). 'Language Deficits, Localization, and Grammar: Evidence for a Distributive Model of Language Breakdown in Aphasic Patients and Neurologically Intact Individuals', *Psychological Review* 108: 759–788.

Diercks, Michael (2012). 'Parameterizing Case: Evidence From Bantu', *Syntax* 15: 253–86.

Diller, Anthony (1993). 'Diglossic Grammaticality in Thai', in William A. Foley (ed.), *The Role of Theory in Language Description*. Berlin: Mouton de Gruyter, 393–420.

Dimmendaal, Gerrit J. (2010). 'Differential Object Marking in Nilo-Saharan', *Journal of African Languages and Linguistics* 31: 13–46.

Dixon, Robert M. W. (1994). *Ergativity*. Cambridge Studies In Linguistics 69. Cambridge: Cambridge University Press.

Dixon, Robert M. W. (1997). *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.

Doctor, Raimond (2004). *A Grammar of Gujarati*. Munich: Lincom Europa.

Dol, Philomena (1999). *A Grammar of Maybrat: A Language of the Bird's Head, Irian Jaya, Indonesia*. Ph.D. dissertation, Leiden University.

Donohue, Mark (2004). *A Grammar of the Skou Language of New Guinea*. MS, University of Singapore. http://pubman.mpdl.mpg.de/pubman/faces/viewItemFullPage.jsp?itemId=escidoc:402710:3 (Accessed 5th March 2014).

Dowty, David R. (1996). 'Towards a Minimalist Theory of Syntactic Structure', in H. Bunt and A. van Horck (eds.), *Discontinuous Constituency*. Berlin: Mouton de Gruyter.

Dresher, Elan (1999). 'Charting the Learning Path: Cues To Parameter Setting', *Linguistic Inquiry* 30: 27–67.

Dressler, Wolfgang U. (1991). 'The Sociolinguistic and Patholinguistic Attrition of Breton Phonology, Morphology, and Morphonology', in H. W. Seliger and R. M. Vago (eds.), *First Language Attrition*. Cambridge: Cambridge University Press.

Dryer, Matthew S. (1992). 'The Greenbergian Word Order Correlations', *Language* 68: 81–138.

Dryer, Matthew S. (1998). 'Aspects of Word Order in the Languages of Europe', in A. Siewierska (ed.), *Constituent Order in the Languages of Europe*. Berlin: Walter De Gruyter, 283–320.

Dryer, Matthew S. (2011). 'Genealogical Language List', in Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. http://wals.info/languoid/genealogy (Accessed 5th March 2014).

Dryer, Matthew S. and Haspelmath, Martin (eds.) (2011). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. http://wals.info/.

Dubenion-Smith, Shannon A. (2010). 'Verbal Complex Phenomena in West Central German: Empirical Domain and Multi-Causal Account', *Journal of Germanic Linguistics* 22: 99–191.

Duffield, Nigel (2001). 'On Certain Head-Final Effects in Vietnamese', in K. Megerdoomian and L.A. Bar-El (eds.), *Proceedings of WCCFL 20*. Somerville, MA: Cascadilla Press, 150–63.

Duncan, Lachlan (2003). 'The Syntactic Structure of Tz'utujil Maya', in Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG03 Conference*. Stanford, CA: CSLI Publications, 164–183.

Dwyer, David (1986). 'What Are Chimpanzees Telling Us About Language?', *Lingua* 69: 219–244.

Eaton, Helen (2003). 'Are There Serial Verb Constructions in Sandawe?'. Paper presented at 7th LASU conference, University of Dar es Salaam, 21–23 August 2003. http://www.drhelenipresume.com/papers/LASU7_paper_Eaton.pdf (Accessed 5th March 2014).

Edwards, John (1994). *Multilingualism*. London: Penguin.

EEBO TCP Corpus: *Early English Books Online - Text Creation Partnership*. http://www.textcreationpartnership.org/tcp-eebo/ (Accessed 5th March 2014).

É.Kiss, Katalin (1995). *Discourse Configurational Languages*. New York: Oxford University Press.

Elbourne, Paul (2005). *Situations and Individuals*. Cambridge, MA: MIT Press.

Elman, Jeffrey (1990). 'Finding Structure in Time', *Cognitive Science* 14: 179–211.

Embick, David, Marantz, Alec, Miyashita, Yasushi, O'Neil, Wayne and Sakai, Kuniyoshi L. (2000). 'A Syntactic Specialization for Broca's Area', *Proceedings of The National Academy of Sciences* 97: 6150–6154.

Emonds, Joseph E. (1976). *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*. New York: Academic Press.

Emonds, Joseph E. (1980). 'Word Order in Generative Grammar', *Journal of Linguistic Research* 1: 33–54.

Enard, Wolfgang, Przeworski, Molly, Fisher, Simon E., Lai, Cecilia S. L., Wiebe, Victor, Kitano, Takashi, Monaco, Anthony P., and Pääbo, Svante (2002). 'Molecular Evolution of FOXP2, a Gene Involved in Speech and Language', *Nature* 418: 869–872.

Enfield, Neil (2003). *Linguistic Epidemiology. Semantics and Grammar of Language Contact in Mainland Southeast Asia.* London: Routledge.

Enfield, Nicholas J. (2008). 'Transmission Biases in Linguistic Epidemiology', *Journal of Language Contact* 2: 299–310.

Engdahl, Elisabet (1982). 'Restrictions on Unbounded Dependencies in Swedish', in E. Engdahl and E. Ejerhed (eds.), *Readings on Unbounded Dependencies in Scandinavian Languages.* Stockholm: Almqvist and Wiksell, 151–174.

Evans, Nicholas, and Levinson, Stephen C. (2009). 'The Myth of Language Universals: Language Diversity and Its Importance for Cognitive Science', *Behavioral and Brain Sciences* 32: 429–448.

Everett, Daniel (1986). 'Pirahã', in Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. 1. Berlin: Mouton de Gruyter, 200–325.

Everett, Daniel (2005). 'Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language', *Current Anthropology* 46: 621–646.

Everett, Daniel (2009). 'Pirahã Culture and Grammar: A Response to Some Criticisms', *Language* 85: 405–442.

Fanselow, Gisbert (2003). 'Surprising Specifiers and Cyclic Spellout', *Generative Linguistics in Poland* 5: 29–46.

Fanselow, Gisbert (2004). 'Cyclic Phonology-Syntax-Interaction: Movement to First Position in German', in S. Ishihara, M. Schmitz, and A. Schwarz (eds.), *Interdisciplinary Studies on Information Structure*. Potsdam: Potsdam University Press, 1–42.

Fanselow, Gisbert, and Lenertová, Denisa (2011). 'Left Peripheral Focus: Mismatches between Syntax and Information Structure', *Natural Language and Linguistic Theory* 29: 169–209.

Fasanella, Adriana and Fortuny, Jordi (2013). 'Deriving Linguistic Variation from Learnability Conditions: The Chunking Procedure'. Unpublished MS, Universitat Autònoma De Barcelona.

Feldman, Heidi, Goldin-Meadow, Susan, and Gleitman, Lila (1978). 'Beyond Herodotus: The Creation of Language by Linguistically Deprived Deaf Children', in A. Lock (ed.), *Action, Symbol, and Gesture: The Emergence of Language*, 350–414.

Fenk-Oczlon, Gertraud, and Fenk, August (1999). 'Cognition, Quantitative Linguistics, and Systemic Typology', *Linguistic Typology* 3: 151–177.

Ferreira, Fernanda, and Patson, Nikole (2007). 'The Good Enough Approach to Language Comprehension', *Language and Linguistics Compass* 1: 71–83.

Fillmore, Charles J., Kay, Paul, and O'Connor, Mary Catherine (1988). 'Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone', *Language* 64: 501–539.

Fischer, Susan (1975). 'Influences on Word Order Change in American Sign Language', in C.N. Li (ed.), *Word Order and Word Order Change*, University of Texas Press, Austin TX, 1–25.

Fischer, Susan (1978). 'Sign Language and Creoles', in P. Siple (ed.), *Understanding Language through Sign Language Research*, New York: Academic Press, 309–331.

Fischer, Susan, and Gong, Qunhu (2010). 'Variation in East Asian Sign Language Structures', in D. Brentari (ed.), *Sign Languages*. Cambridge: Cambridge University Press, 499–518.

Fitch, W. Tecumseh, and Friederici, Angela D. (2012). 'Artificial Grammar Learning Meets Formal Language Theory: An Overview', *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1598): 1933–1955.

Fitch, W. Tecumseh, Hauser, Marc D., and Chomsky, Noam (2005). 'The Evolution of the Language Faculty: Clarifications and Implications', *Cognition* 97: 179–210.

Fodor, Janet D., (1978). 'Parsing Strategies and Constraints on Transformations', *Linguistic Inquiry* 9: 427–473.

Fodor, Jerry A., Bever, Tom G., and Garrett, Merrill F. (1974). *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw Hill.

Foley, William A. (1991). *The Yimas Language of New Guinea*. Stanford, CA: Stanford University Press.

Fortescue, Michael (1984). *West Greenlandic*. London: Croom Helm.

Fortson, Benjamin W. (1997). *Indo-European Language and Culture: An Introduction*. Oxford: Blackwell.

Fortuny, Jordi (2008). *The Emergence of Order in Syntax*. Amsterdam: John Benjamins.

Fortuny, Jordi (2010). 'On the Duality of Patterning', in M. De Vries and J-W. Zwart (eds.), *Structure Preserved. Studies in Syntax for Jan Koster*. Amsterdam: Benjamins, 131–140.

Francis, Elaine, and Michaelis-Cummings, Laura (2012). 'Effects of Weight and Definiteness on Speakers' Choice of Clausal Ordering in English'. Paper presented at the Annual Meeting of the Linguistic Society of America, Portland, OR, January 2012.

Frank, Paul (1990). *Ika Syntax*. Arlington, TX: Summer Institute of Linguistics and The University of Texas.

Franks, Steven (1995). *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.

Frazier, Lyn (1985). 'Syntactic Complexity', in D. Dowty, L. Karttunen, and A. Zwicky (eds.), *Natural Language Parsing*. Cambridge: Cambridge University Press, 129–189.

Frazier, Lyn (1987). 'Sentence Processing: A Tutorial Review', in Max Coltheart (ed.), *Attention and Performance XII*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Friederici, Angela D. (2011). 'The Brain Basis of Language Processing: From Structure to Function', *Physiological Reviews* 91: 1357–1392.

Friederici, Angela D., Fiebach, Christian J., Schlesewsky, Matthias, Bornkessel, Ina D., and von Cramon, D. Yves (2006). 'Processing Linguistic Complexity and Grammaticality in the Left Frontal Cortex', *Cerebral Cortex* 16: 1709–1717.

Friedmann, Na'ama (2006). 'Speech Production In Broca's Agrammatic Aphasia: Syntactic Tree Pruning', in Y. Grodzinsky and K. Amunts (eds.), *Broca's Region*. New York: Oxford University Press, 63–82.

Friedmann, Na'ama, Belletti, Adriana and Rizzi, Luigi (2009). 'Relativized Relatives: Types of Intervention in the Acquisition of A-Bar Dependencies', *Lingua* 119: 67–88.

Friedmann, Na'ama, and Grodzinsky, Yosef (1997). 'Tense and Agreement In Agrammatic Production: Pruning The Syntactic Tree', *Brain and Language* 56: 397–425.

Friedmann, Na'ama, and Shapiro, Lewis P. (2003). 'Agrammatic Comprehension of Simple Active Sentences With Moved Constituents: Hebrew OSV and OVS Structures', *Journal of Speech, Language and Hearing* 46: 288–297.

Fromkin, Victoria (1993). 'Speech Production', in J.B. Gleason and N.B. Ratner (eds.), *Psycholinguistics*. Orlando, FL: Harcourt Brace Jovanovich.

Fromkin, Victoria, and Rodman, Robert (1983). *An Introduction to Language* (3rd edition). New York: Holt, Rinehart and Winston.

Fromkin, Victoria, Rodman, Robert, and Hyams, Nina (2010). *An Introduction to Language* (9th edition). Boston: Wadsworth.

Fuller, Judith Wheaton (1985). *Topic and Comment in Hmong*. Ph.D. dissertation, University of Minnesota.

Gahl, Susanne, Menn, Lise, Ramsberger, Gail, Jurafsky, Daniel S., Elder, Elizabeth, Rewega, Molly, and Holland, Audrey L. (2003). 'Syntactic Frame and Verb Bias in Aphasia: Plausibility Judgments of Undergoer-Subject Sentences', *Brain and Cognition* 2003, 53: 223–228.

Gallego, Angel (2013). 'Object Shift in Romance', *Natural Language and Linguistic Theory* 31: 409–451.

Ganong, W. Francis, III. (1980). 'Phonetic Categorization in Auditory Word Perception', *Journal of Experimental Psychology: Human Perception and Performance* 6: 110–125.

Gee, James Paul, and Goodhart, Wendy (1985). 'Nativization, Linguistic Theory and Deaf Language Acquisition', *Sign Language Studies* 49: 291–342.

Gell-Mann, Murray (1994). *The Quark and the Jaguar: Adventures in the Simple and the Complex*. New York: W.H. Freeman and Company.

Gell-Mann, Murray (1995). 'What Is Complexity?', *Complexity* 1: 16–19.

Gennari, Silvia, and MacDonald, Maryellen C. (2009). 'Linking Production and Comprehension Processes: The Case of Relative Clauses', *Cognition* 111: 1–23.

Gennari, Sylvia, and Poeppel, David (2003). 'Processing Correlates of Lexical Semantic Complexity', *Cognition* 89: B27–B41.

Gibson, Edward (1998). 'Linguistic Complexity: Locality of Syntactic Dependencies', *Cognition* 68: 1–76.

Gibson, Edward (2000). 'The Dependency Locality Theory: A Distance-based Theory of Linguistic Complexity', in A. Marantz, Y. Miyashita, and W. O'Neil (eds.), *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*. Cambridge, MA: MIT Press, 95–126.

Gibson, Edward, and Thomas, James (1999). 'Memory Limitations and Structural Forgetting: The Perception of Complex Ungrammatical Sentences as Grammatical', *Language and Cognitive Processes* 14: 225–248.

Gil, David (2001a). 'Creoles, Complexity and Riau Indonesian', *Linguistic Typology* 5: 325–371.

Gil, David (2001b). 'Escaping Eurocentrism: Fieldwork as a Process of Unlearning', in P. Newman and M. Ratliff (eds.), *Linguistic Fieldwork*. Cambridge: Cambridge University Press, 102–132.

Gil, David (2005a). 'Isolating-Monocategorial-Associational Language', in H. Cohen and C. Lefebvre (eds.), *Categorization in Cognitive Science*, Elsevier, Oxford, 347–379.

Gil, David (2005b). 'Word Order Without Syntactic Categories: How Riau Indonesian Does It', in A. Carnie, H. Harley and S.A. Dooley (eds.), *Verb First: On the Syntax of Verb-Initial Languages*, Amsterdam: John Benjamins, 243–263.

Gil, David (2006). 'Intonation and Thematic Roles in Riau Indonesian', in C.M. Lee, M. Gordon, and D. Büring (eds.), *Topic and Focus, Cross-Linguistic Perspectives on Meaning and Intonation*, Studies in Linguistics and Philosophy 82. Dordrecht: Springer, 41–68.

Gil, David (2007). 'Creoles, Complexity and Associational Semantics', in Umberto Ansaldo, Stephen Matthews, and Lisa Lim (eds.), *Deconstructing Creole: New Horizons in Language Creation*. Amsterdam: John Benjamins, 67–108.

Gil, David (2008). 'How Complex Are Isolating Languages?', in Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 109–131.

Gil, David (2009). 'How Much Grammar Does it Take to Sail a Boat?', in Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 19–33.

Gil, David (2012a). 'How Did Predication Evolve?', in T.C. Scott-Phillips, M. Tamariz, E.A. Cartmill and J.R. Hurford (eds.), *The Evolution of Language, Proceedings of the 9th International Conference (EVOLANG9)*. World Scientific Publishing, New Jersey, 134–141.

Gil, David (2012b). 'Where Does Predication Come From?', *Canadian Journal of Linguistics* 57: 303–333.

Gil, David (in progress). 'Null Content Interrogatives', in M.S. Dryer and M. Haspelmath (eds.), *World Atlas of Language Structures Online,* Max Planck Digital Library, Munich. http://www.wals.info/.

Gil, David (to appear). 'Tense, Aspect and Mood Marking', in M.S. Dryer and M. Haspelmath (eds.), *World Atlas of Language Structures Online,* Max Planck Digital Library, Munich. http://www.wals.info/.

Gillon, Carrie (2006). *The Semantics of Determiners: Domain Restriction in Sḵwx̱wú7mesh*. Ph. D. dissertation, University of British Columbia.

Giorgi, Alessandra, and Pianesi, Fabio (1997). *Tense and Aspect. From Semantics to Morphosyntax.* Oxford: Oxford University Press.

Givón, Talmy (1979). *On Understanding Grammar*. New York: Academic Press.

Givón, Talmy (1981). 'Typology and Functional Domains', *Studies in Language* 5(2): 163–193.

Givón, Talmy (2009). *The Genesis of Syntactic Complexity: Diachrony, Ontogeny, Neurocognition, Evolution.* Amsterdam: John Benjamins.

Gleitman, Lila R. (1965). 'Coordinating Conjunctions in English', *Language* 41: 260–293.

Göksel, Aslı, and Kerslake, Celia (2005). *Turkish: A Comprehensive Grammar*. London: Routledge.

Goldberg, Adele E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Goldberg, Adele E. (1996). 'Jackendoff and Construction-based Grammar', *Cognitive Linguistics* 7: 3–19.

Goldberg, Adele E. (2006). *Constructions at Work*. New York: Oxford University Press.

Goldin-Meadow, Susan (2003). *The Resilience of Language*. Psychology Press.

Goldin-Meadow, Susan (2005). 'What Language Creation in the Manual Modality Tells us about the Foundations of Language', *The Linguistic Review* 22: 199–225.

Gonda, Jan (1956). *The Character of the Indo-European Moods (With Special Regard to Greek and Sanskrit)*. Weisbaden: Harrassowitz.

Goodglass, Harold, Kaplan, Edith, and Barresi, Barbara (2001). *Boston Diagnostic Aphasia Examination*, (3rd edition). Austin TX: Pro-Ed.

Gordon, Peter C., Hendrick, Randall, and Johnson, Marcus (2001). 'Memory Interference during Language Processing', *Journal of Experimental Psychology: Learning, Memory and Cognition* 27: 1411–1423.

Gordon, Peter C., Hendrick, Randall, and Johnson, Marcus (2004). 'Effects of Noun Phrase Type on Sentence Complexity', *Journal of Memory and Language* 51: 97–114.

Gordon, Peter C., Hendrick, Randall, Johnson, Marcus, and Lee, Yoonhyoung (2006). 'Similarity-based Interference during Language Comprehension: Evidence From Eye Tracking During Reading', *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32: 1304–1321.

Gordon, Raymond G. (ed.) (2005). *Ethnologue: Languages of the World, Fifteenth Edition*. Dallas, TX: Summer Institute of Linguistics.

Greenberg, Joseph H. (1960). 'A Quantitative Approach to the Morphological Typology of Language', *International Journal of American Linguistics* 26: 192–220.

Greenberg, Joseph H. (1963). 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements', in J.H. Greenberg (ed.), *Universals of Language*. Cambridge, MA: MIT Press, 73–113.

Greenberg, Joseph H., Ferguson, Charles A., and Moravcsi, Edith (eds.) (1978). *Universals of Human Language, Volume 2: Phonology*. Stanford, CA: Stanford University Press.

Greenfield, Patricia Marks, and Savage-Rumbaugh, Sue (1990). 'Grammatical Combination in Pan Paniscus: Processes of Learning and Invention in the Evolution and Development of Language', in S. T. Parker and K. R. Gibson (eds.), *'Language' and Intelligence in Monkeys and Apes, Comparative Developmental Perspectives*. Cambridge: Cambridge University Press, 540–578.

Grodner, Daniel J., and Gibson, Edward A. F. (2005). 'Consequences of the Serial Nature of Linguistic Input for Sentential Complexity', *Cognitive Science* 29: 261–291.

Grodzinsky, Yosef (2000). 'The Neurology of Syntax: Language Use Without Broca's Area', *Behavioral and Brain Sciences* 23: 1–21.

Grodzinsky, Yosef (2006). 'A Blueprint for a Brain Map of Syntax', in Y. Grodzinsky and K. Amunts (eds.), *Broca's Region*. Oxford: Oxford University Press, 83–107.

Grodzinsky, Yosef (2010). 'The Picture of The Linguistic Brain: How Sharp Can It Be? Reply to Fedorenko and Kanwisher', *Language and Linguistics Compass* 4: 605–622.

Grodzindky, Yosef, and Friederici, Angela D. (2006). 'Neuroimaging of Syntax and Syntactic Processing', *Current Opinion in Neurobiology* 16: 240–246.

Grodzinsky, Yosef, Wexler, Kenneth, Chien, Yu-Chin, Marakovitz, Susan, and Solomon, Julie (1993). 'The Breakdown of Binding Relations', *Brain and Language* 45: 396–422.

Guasti, Maria Teresa (2002). *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press.

Guirardello, Raquel (1999). *A Reference Grammar of Trumai*. Ph.D. dissertation, Rice University.

Guirardello-Damien, Raquel (2010). 'Ergativity in Trumai', in Spike Gildea and Francesc Queixalós (eds.), *Ergativity in Amazonia*. Amsterdam: John Benjamins, 203–234.

Hackl, Martin (2009). 'On The Grammar and Processing of Proportional Quantifiers: Most Versus More Than Half', *Natural Language Semantics* 17: 63–98.

Haegeman, Liliane (1994). 'Verb Raising as Verb Projection Raising: Some Empirical Problems', *Linguistic Inquiry*, 25: 509–522.

Haegeman, Liliane, and van Riemsdijk, Henk (1986). 'Verb Projection Raising, Scope and the Typology of Rules Affecting Verbs', *Linguistic Inquiry* 17: 417–66.

Hafri, Alon, Papafragou, Anna, and Trueswell, Jon (2013). 'Getting the Gist of Events: Recognition of Two-participant Actions from Brief Displays', *Journal of Experimental Psychology: General*, 142(3): 880–905.

Hagman, Roy Stephen (1977). *Nama Hottentot Grammar*. Bloomington, IN: Indiana University.

Haider, Hubert (1995). 'Downward to the Right', in Uli Lutz and Jürgen Pafel (eds.), *On Extraction and Extraposition in German*. Amsterdam: John Benjamins, 145–271.

Haider, Hubert (1997). 'Economy in Syntax is Projective Economy', in Chris Wilder, Hans-Martin Gärtner, and Manfred Bierwisch (eds.), *The Role of Economy Principles in Linguistic Theory*. Berlin: Academie-Verlag, 205–226.

Haider, Hubert (2000). 'OV is More Basic Than VO', in Peter Svenonius (ed.), *The Derivation of VO and OV*. Amsterdam: John Benjamins, 45–67.

Haider, Hubert (2003). 'V-Clustering and Clause Union', in Pieter A.M. Seuren and Gerard Kempen (eds.), *Verb Constructions in German and Dutch*. Amsterdam: John Benjamins, 91–126.

Haldane, John Burdon Sanderson (1927). 'A Mathematical Theory of Natural and Artificial Selection. Part V. Selection and Mutation', *Proceedings of the Cambridge Philosophical Society* 23: 838–844.

Hale, John T. (2001). 'A Probabilistic Early Parser as a Psycholinguistic Model', *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies*, 1–8.

Hale, John T. (2003). 'The Information Conveyed By Words in Sentences', *Journal of Psycholinguistic Research* 32: 101–123.

Hale, Kenneth (1970). 'The Passive and Ergative In Language Change: The Australian Case', in S. Wurm and D. Laysock (eds.), *Pacific Linguistic Studies In Honor of Arthur Capell*. Pacific Linguistic Series. Canberra: Linguistics Circle of Canberra, 757–781.

Hale, Kenneth, and Keyser, S. Jay (2002). *Prolegomena to a Theory of Argument Structure*. Linguistic Inquiry Monograph 39. Cambridge, MA: MIT Press.

Hale, Mark (1987). 'Notes on Wackernagel's Law in the Language of the Rigveda', in C. Watkins (ed.), *Studies In Memory of Warren Cowgill*. Berlin: Mouton De Gruyter, 38–50.

Hardman, Martha J. (1966). *Jaqaru: Outline of Phonological and Morphological Structure*. The Hague: Mouton.

Hardman, Martha J. (2000). *Jaqaru*. Munich: LINCOM Europa.

Hargreaves, Katie (2004). '*to* and *and* Complementation in English'. Unpublished MA thesis, University of Cambridge.

Hargreaves, Katie (2005). '*to*- and *and*-complementation in English', *Durham and Newcastle Working Papers in Linguistics* 11: 27–40.

Harrell, Frank E. Jr. (2012). *RMS: Regression Modeling Strategies*. R package version 3.5-0. http://CRAN.R-project.org/package=rms (Accessed 5th March 2014).

Harris, Alice C. (1981). *Georgian Syntax: A Study in Relational Grammar*. Cambridge: Cambridge University Press.

Harris, Alice C., and Campbell, Lyle (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin (2008). 'Frequency vs. Iconicity in Explaining Grammatical Asymmetries', *Cognitive Linguistics* 19: 1–33.

Haspelmath, Martin, Dryer, Matthew S., Gil, David, and Comrie, Bernard (eds.) (2005). *World Atlas of Language Structures*, Oxford: Oxford University Press.

Haspelmath, Martin, Dryer, Matthew S., Gil, David, and Comrie, Bernard (2008). *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. http: //wals.info/.

Haudricourt, André G. (1961). 'Richesse en Phonèmes et Richesse en Locuteurs', *L'Homme* 1: 5–10.

Hauser, Marc D., Chomsky, Noam, and Fitch, W. Tecumseh (2002). 'The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?', *Science* 298: 1569–1579.

Hawkins, John A. (1982). 'Cross-Category Harmony, X-Bar and the Predictions of Markedness', *Journal of Linguistics* 18: 1–35.

Hawkins, John A. (1983). *Word Order Universals*. New York: Academic Press.

Hawkins, John A. (1986). A *Comparative Typology of English and German: Unifying the Contrasts*. Austin: University of Texas Press.

Hawkins, John A. (1994). A *Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. (1999). 'Processing Complexity and Filler Gap Dependencies Across Grammars', *Language*, 75: 244–85.

Hawkins, John A. (2000). 'The Relative Ordering of Prepositional Phrases in English: Going Beyond Manner-Place-Time', *Language Variation and Change* 11: 231–266.

Hawkins, John A. (2001). 'Why Are Categories Adjacent?', *Journal of Linguistics* 37: 1–34.

Hawkins, John A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Hawkins, John A. (2006). 'Gradedness as Relative Efficiency in the Processing of Syntax and Semantics', in Gisbert Fanselow, Caroline Féry, Ralf Vogel and Mathias Schlesewsky (eds.), *Gradience in Grammar*. Oxford: Oxford University Press.

Hawkins, John A. (2009). 'An Efficiency Theory of Complexity and Related Phenomena', in G. Sampson, D. Gil and P. Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 252–268.

Hawkins, John A. (2014). *Cross-Linguistic Variation and Efficiency*. Oxford: Oxford University Press.

Hay, Jennifer, and Bauer, Laurie (2007). 'Phoneme Inventory Size and Population Size', *Language* 83: 388–400.

Hayes, Bruce (2009). *Introductory Phonology*. Oxford: Blackwell.

Headland, Thomas N. (1974). 'Grammatical Sketch of Casiguran (Dumagat)', in T.N. Headland, A. Healey and J. Witucki (eds.), *Papers in Philippine Linguistics* No. 6, Pacific Linguistics, Series A - No. 43, Canberra, 1–54.

Heim, Irene, and Kratzer, Angelika (1998). *Semantics and Generative Grammar*. Oxford: Blackwell.

Heine, Bernd, and Kuteva, Tania (2007). *The Genesis of Grammar. A Reconstruction*. Oxford: Oxford University Press.

Henderson, James (1995). *Phonology and Grammar of Yele, Papua New Guinea*. Canberra: Australian National University.

Hendery, Rachel (2012). *Relative Clauses in Time and Space. A Case Study in the Methods of Diachronic Typology*. Amsterdam: John Benjamins.

Hendriks, Bernadet (2008). 'I have Therefore I Exist: Possession in Jordanian Sign Language', in U. Zeshan and P. Perniss (eds.), *Possessive and Existential Constructions in Sign Languages*, 55–70.

Heycock, Caroline, and Petersen, Hjalmar P. (2012). 'Pseudo-coordination in Faroese', in K. Braunmueller, and C. Gabriel (eds.), *Multilingual Individuals and Multilingual Societies*. Hamburg: John Benjamins B.V., 259–280.

Hillis, Argye E., and Caramazza, Alfonso (1991). 'Category-Specific Naming and Comprehension Impairment: A Double Dissociation', *Brain* 114: 2081–2094.

Hinds, John (1989). 'Left-Dislocation in Thai', *Language and Linguistics* 2: 61–83.

Hinton, Geoffrey E., and Nowlan, Steven J. (1987). 'How Learning Can Guide Evolution', *Complex Systems* 1: 495–502.

Hockett, Charles F. (1955). *A Manual of Phonology*. Baltimore, MD: Waverly Press.

Hockett, Charles F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.

Hockett, Charles F. (1963). 'The Problem of Universals in Language', in Joseph H. Greenberg (ed.), *Universals of Language*. Cambridge, MA: MIT Press, 1–29.

Hofmeister, Philip, and Sag, Ivan A. (2010). 'Cognitive Constraints and Island Effects', *Language* 86: 366–415.

Holmberg, Anders (2010). 'Null Subject Parameters', in T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation*. Cambridge: Cambridge University Press, 88–124.

Holmberg, Anders, and Platzack, Christer (1995). *The Role of Inflection in Scandinavian Syntax*. Oxford: Oxford University Press.

Holmberg, Anders, and Roberts, Ian (2013). 'The Syntax-Morphology Relation', *Lingua* 130: 111–131.

Holmberg, Anders, Nayudu, Aarti, and Sheehan, Michelle (2009). 'Three Partial Null-Subject Languages: A Comparison of Brazilian Portuguese, Finnish and Marathi', *Studia Linguistica* 63: 59–97.

Hommerberg, Charlotte, and Tottie, Gunnel (2007). '*Try to* or *try and*? Verb Complementation in British and American English', *ICAME Journal* 31: 45–64.

Hopcroft, John E., Motwani, Rajeev and Ullman, Jeffrey D. (2001). *Introduction to Automata Theory, Languages and Computation* (2nd Edition). Boston: Addison-Wesley.

Hornstein, Norbert, and Pietroski, Paul (2009). 'Basic Operations: Minimal Syntax-Semantics', *Catalan Journal of Linguistics* 8: 113–139.

Hornstein, Norbert, and Weinberg, Amy (1981). 'Case Theory and Preposition Stranding', *Linguistic Inquiry* 12: 55–92.

Hornstein, Norbert, Nunes, Jairo, and Grohmann, Kleanthes K. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.

Huang, C.-T. James (1982a). *Logical Relations in Chinese and the Theory of Grammar*. Ph. D. dissertation, MIT.

Huang, C.-T. James (1982b). 'Move-*Wh* in a Language Without *Wh*-Movement', *The Linguistic Review* 1: 369–416.

Huang, C.-T. James (1984). 'On the Distribution and Reference of Empty Pronouns', *Linguistic Inquiry* 15: 531–574.

Huang, C.-T. James (2013). 'On Syntactic Analyticity and Parametric Theory'. To appear in C.-T. James Huang, Andrew Simpson and Audrey Li (eds.), *Handbook of Chinese Linguistics*. Oxford: Wiley-Blackwell.

Huang, C.-T. James, Li, Y.-H. Audrey, and Li, Yafei (2009). *The Syntax of Chinese*. Cambridge: Cambridge University Press.

Hübler, Alfred W. (2007). 'Understanding Complex Systems', *Complexity* 12: 9–11.

Hunsicker, Dea, and Goldin-Meadow, Susan (2012). 'Hierarchical Structure in a Self-Created Communication System: Building Nominal Constituents in Homesign', *Language* 88: 732–763.

Hurford, James R. (2007). *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.

Hurford, James (2012). *The Origins of Grammar: Language in the Light of Evolution*. Oxford: Oxford University Press.

Ikeda, Edward (1991). *Sentential Complementation in Mohawk*. MA thesis: McGill University. http://www.collectionscanada.gc.ca/obj/thesescanada/vol2/QMM/TC-QMM-61093.pdf (Accessed 4th March 2014).

Indefrey, P., Hagoort, P., Herzog, H., Seitz, J., and Brown, C.M. (2001). 'Syntactic Processing in Left Prefrontoal Cortex is Independent of Lexical Meaning', *Neuroimage* 14: 546–555.

Iwasaki, Shoichi, and Ingkaphirom, Preeya (2005). *A Reference Grammar of Thai*. Cambridge: Cambridge University Press.

Jackendoff, Ray (1977). *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1987). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1999). 'Possible Stages in the Evolution of the Language Capacity', *Trends in Cognitive Sciences* 3: 272–279.

Jackendoff, Ray (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, Ray (2003). 'Précis of *Foundations of Language: Brain, Meaning, Grammar, Evolution*', *Behavioral and Brain Sciences* 26: 651–665.

Jackendoff, Ray (2007). 'A Parallel Architecture Perspective on Language Processing', *Brain Research* 1146: 2–22.

Jackendoff, Ray (2008). 'Alternative Minimalist Visions of Language', *Chicago Linguistic Society* 41: 189–226.

Jackendoff, Ray (in preparation). *What Can You Say Without Syntax*?

Jaeger, T. Florian, and Tily, Harry (2011). 'Language Processing Complexity and Communicative Efficiency', *WIREs: Cognitive Science* 2: 323–335.

Jakobson, Roman (1971 [1929]). 'Remarques sur l'Evolution Phonologique du Russe Comparée à Celle des autres Langues Slaves', in Roman Jakobson (ed.), *Selected Writings, Vol. 1: Phonological Studies*. The Hague: Mouton, 6–116.

Jakobson Roman (1956). 'Two Aspects of Language and Two Types of Aphasic Disturbances', in Roman Jakobson and Morris Halle, *Fundamentals of Language*. The Hague: Mouton.

Janssen, Dirk, Bickel, Balthasar, and Zúñiga, Fernando (2006). 'Randomization Tests in Language Typology', *Linguistic Typology* 10: 419–440.

Jelinek, Eloise (1984). 'Empty Categories, Case, and Configurationality', *Natural Language and Linguistic Theory* 2: 39–76.

Johnson-Laird, Philip N., and Stevenson, Rosemary (1970). 'Memory for Syntax', *Nature* 227: 412.

Johnston, Trevor, and Schembri, Adam (2007). *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.

Joseph, John E., and Newmeyer, Frederick J. (2012). 'All Languages are Equally Complex: The Rise and Fall of a Consensus', *Historiographia Linguistica* 39: 341–368.

Joshi, Aravind K. (1985). 'How Much Context-Sensitivity is Necessary for Characterizing Structural Descriptions?', in D. Dowty, L. Karttunen, and A. Zwicky (eds.), *Natural Language Processing: Theoretical, Computational, and Psychological Perspectives*. New York, NY: Cambridge University Press, 206–250.

Julien, Marit (2002). *Syntactic Heads and Word Formation*: New York: Oxford University Press.

Juola, Patrick (1998). 'Measuring Linguistic Complexity: The Morphological Tier', *Journal of Quantitative Linguistics* 5: 206–213.

Juola, Patrick (2008). 'Assessing Linguistic Complexity', in Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 89–108.

Just, Marcel A., Carpenter, Patricia A., Keller, Timothy A., Eddy, William F., and Thulborn, Keith R. (1996). 'Brain Activation Modulated by Sentence Comprehension'. *Science* 274: 114–116.

Justeson, John S., and Stephens, Laurence D. (1984). 'On the Relationship Between the Numbers of Vowels and Consonants in Phonological Systems', *Linguistics* 22: 531–545.

Kabak, Bans (2004). 'Acquiring Phonology is not Acquiring Inventories but Contrasts: The Loss of Turkic and Korean Primary Long Vowels', *Linguistic Typology* 8: 351–368.

Kakumasu, James (1986). 'Urubu-Kaapor', in Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. 1. Berlin: Mouton de Gruyter, 326–403.

Kandel, Eric R., Schwartz, James H., Jessell, Thomas M., Siegelbaum, Steven A., and Hudspeth, A. J. (2012). *Principles of Neural Science* (5th edition). New York: McGraw Hill Medical.

Karlsson, Fred (2007). 'Constraints on multiple center-embedding of clauses', *Journal of Linguistics* 43: 365–392.

Kathol, Andreas (2000). *Linear Syntax*. Oxford: Oxford University Press.

Katz, Jerrold J. (1976). 'A Hypothesis about the Uniqueness of Natural Language', in Stevan R. Harnad, Horst D. Steklis, and Jane Lancaster (eds.), *Origins and Evolution of Language and Speech*. New York: New York Academy of Sciences, 33–41.

Kawasha, Boniface (2003). *Lunda Grammar: A Morphosyntactic and Semantic Analysis*. Ph. D. dissertation, University of Oregon.

Kay, Paul, and Fillmore, Charles J. (1999). 'Grammatical Constructions and Linguistic Generalizations: The What's X Doing Y? Construction', *Language* 75: 1–33.

Kayne, Richard (1981). 'On Certain Differences Between French and English', *Linguistic Inquiry* 12: 349–371.

Kayne, Richard (1989). 'Notes on English Agreement', *CIEFL Bulletin, Hyderabad Working Papers in Linguistics* 1: 41–67.

Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Keenan, Edward L., and Comrie, Bernard (1977). 'Noun Phrase Accessibility and Universal Grammar', *Linguistic Inquiry* 8: 63–99.

Keenan, Edward L., and Paperno, Denis (2012). *Handbook of Quantifiers in Natural Language.* Dordrecht: Springer.

Kegl, Judy, Senghas, Ann, and Coppola, Marie (1999). 'Creation through Contact: Sign Language Change and Sign Language Emergence in Nicaragua', in M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge MA, 179–237.

Kemenade, Ans van (1987). *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris Publications.

Kemmer, Suzanne (1994). 'Middle Voice, Transitivity, and the Elaboration of Events', in B. Fox and P. J. Hopper (eds.), *Voice: Form and Function*. Amsterdam/Philadelphia: John Benjamins, 179–230.

Kenesei, István, Vago, Robert M., and Fenyvesi, Anna (1998). *Hungarian*. London: Routledge.

Kilian-Hatz, Christa (2003). *Khwe Dictionary*: *With a Supplement on Khwe Place-names of West Caprivi by Matthias Brenzinger*. Cologne: Rüdiger Köppe Verlag.

King, Jonathan, and Just, Marcel A. (1991). 'Individual Differences in Syntactic Processing: The Role of Working Memory', *Journal of Memory and Language* 30: 580–602.

Kiparsky, Paul (1968). 'Tense and Mood in Indo-European Syntax', *Foundations of Language* 4: 30–57.

Kiparsky, Paul (1995). 'Indo-European Origins of Germanic Syntax', in A. Battye and I. Roberts (eds.), *Clause Structure and Language Change*. Oxford Studies In Comparative Syntax. Oxford: Oxford University Press, 140–169.

Kiparsky, Paul (1997). 'The Rise of Positional Licensing', in Ans van Kemenade and Nigel Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 460–494.

Kitagawa, Yoshihisa (1986). *Subjects in English and Japanese*. Ph.D. dissertation, University of Massachusetts, Amherst.

Klein, Wolfgang, and Perdue, Clive (1997). 'The Basic Variety (or: Couldn't Natural Languages be Much Simpler?)', *Second Language Research* 13: 301–347.

Kluender, Robert (1998). 'On the Distinction Between Strong and Weak Islands: A Processing Perspective', in Peter W. Culicover and Louise McNally (eds.), *The Limits of Syntax*. New York: Academic Press, 241–279.

Köhler, Reinhard (2007). 'Quantitative Analysis of Syntactic Structures in the Framework of Synergetic Linguistics', in Alexander Mehler and Reinhard Köhler (eds.), *Aspects of Automatic Text Analysis (Studies in Fuzziness and Soft Computing (StudFuzz) 209)*. Berlin: Springer, 191–209.

Kolk, Herman (2006). 'How Language Adapts to The Brain: An Analysis of Agrammatic Aphasia', in L. Progovac, K. Paesani, E. Casielles, and E. Barton (eds.), *The Syntax of Nonsententials: Multidisciplinary Perspectives*. Amsterdam: Benjamins, 229–258.

Kolmogorov, Andrey N. (1965). 'Three Approaches to the Quantitative Definition of Information', reprinted in 1968 in *International Journal of Computer Mathematics* 2: 157–168.

König, Ekkehard, and Gast, Volker (2007). *Understanding English-German Contrasts*. Berlin: Erich Schmidt Verlag.

Koopman, Hilda, and Sportiche, Dominique (1991). 'The Position of Subjects', *Lingua* 85: 211–258.

Kouwenberg, Silvia (1994). *A Grammar of Berbice Dutch Creole*. Berlin: Mouton de Gruyter.

Kramer, Ruth (2010). 'The Amharic Definite Marker and the Syntax-Morphology Interface', *Syntax* 13: 196–240.

Kratzer, Angelika (1981). 'The Notional Category of Modality', in Hans-Jürgen Eikmeyer and Hannes Rieser (eds.), *Words, Worlds, and Contexts. New Approaches in Word Semantics*. Berlin: de Gruyter, 38–74.

Kratzer, Angelika (1998). 'More Structural Analogies between Pronouns and Tenses', in *Proceedings of SALT VIII*, 92–110.

Kroch, Anthony S. (1974). *The Semantics of Scope in English*. Ph.D. dissertation, MIT.

Kroch, Anthony S. (2001). 'Syntactic Change', in M. Baltin and C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, Malden, MA: Blackwell, 699–729.

Kruspe, Nicole D. (1999). *Semelai*. Ph.D. dissertation, University of Melbourne.

Kuperberg, Gina (2007). 'Neural Mechanisms of Language Comprehension: Challenges to Syntax', *Brain Research* 1146: 23–49.

Kuperman, Victor, Schreuder, Robert, Bertram, Raymond, and Baayen, R. Harald (2009). 'Reading Polymorphemic Dutch Compounds: Towards a Multiple Route Model of Lexical Processing', *Journal of Experimental Psychology: Human Perception and Performance* 35: 879–895.

Kuryłowicz, Jerzy (1964). *The Inflectional Categories of Indo-European*. Heidelberg: Carl Winter Universitätsverlag.

Kusters, Wouter (2003). *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Utrecht: LOT.

Kusters, Wouter (2008). 'Complexity in Linguistic Theory, Language Learning and Language Change', in Matti Miestamo, Kaius Sinnemäki and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 3–22.

Kusters, Wouter, and Muysken, Pieter (2001). 'The Complexities of Arguing About Complexity', *Linguistic Typology* 5: 182–185.

Kutsch Lojenga, Constance (1994). *Ngiti: A Central-Sudanic Language of Zaire*. Köln: Köppe.

Kvam, Sigmund (1983). *Linksverschachtelung im Deutschen und Norwegischen*. Tübingen: Max Niemeyer Verlag.

Kwon, Nayoung (2004). 'Syntactic and Semantic Mismatches in the Korean *ko*-construction', in V. Chand, A. Kelleher, A. J. Rodríguez, and B. Schmeiser (eds.), *WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics*. Somerville: Cascadilla Press, 102–114.

Lai, Vicky T., Curran, Tim, and Menn, Lise (2009). 'Comprehending Conventional and Novel Metaphors: An ERP study', *Brain Research* 1284: 145–155.

Laka, Itziar (2006). 'On the Nature of Case in Basque: Structural Or Inherent?', in H. Broekhuis, N. Corver, J. Koster, R. Huybregts and U. Kleinhenz (eds.), *Organizing Grammar: Linguistic Studies in Honor of Henk Van Riemsdijk*. Berlin/New York: Mouton De Gruyter, 374–382.

Lakoff, George (1986). 'Frame Semantic Control of the Coordinate Structure Constraint', in W. H. Eilfort, P. D. Kroeber, and K. L. Peterson (eds.), in *Proceedings of the 21st Annual Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press, 152–168.

Laman, Karl E. (1912). *Grammar of the Kongo Language*. New York: Christian Alliance.

Landman, Meredith (2006). *Variables in Natural Language*. Ph.D. dissertation, University of Massachusetts, Amherst.

Lapointe, Steven G. (1985). 'A Theory of Verb Form Use in the Speech of Agrammatic Aphasics', *Brain and Language* 24: 100–155.

LaPolla, Randy J., and Huang, Chenglong (2003). *A Grammar of Qiang: With Annotated Texts and Glossary*. Berlin: Mouton de Gruyter.

Lardiere, Donna (2008). 'Feature-Assembly in Second Language Acquisition', in Juana Liceras, Helmut Zobl, and Helen Goodluck (eds.), *The Role of Formal Features in Second Language Acquisition*. New York: Lawrence Erlbaum Associates, 106–140.

Lardiere, Donna (2009). 'Some Thoughts on the Contrastive Analysis of Features in Second Language Acquisition', *Second Language Research,* 25: 173–227.

Larson, Richard K. (1988). 'On The Double Object Construction', *Linguistic Inquiry* 19: 335–391.

Lasnik, Howard (2005). 'Grammar, Levels, and Biology', in J. McGilvray (ed.), *The Cambridge Companion to Chomsky*. Cambridge: Cambridge University Press, 60–83.

Lebeaux, David (1989). *Language Acquisition and the Form of the Grammar*. Ph.D. dissertation, University of Massachusetts, Amherst.

Ledgeway, Adam (2012). *From Latin to Romance: Morphosyntactic Typology and Change*. Oxford: Oxford University Press.

Legate, Judy (2008). 'Morphological and Abstract Case', *Linguistic Inquiry* 39: 55–101.

Legate, Judy (2011). 'Types of Ergativity', *Lingua* 122: 181–191.

Lehmann, Christian (1985a). 'Ergative and Active Traits in Latin', in F. Plank (ed.), *Relational Typology*. Amsterdam: Mouton, 243–255.

Lehmann, Christian (1985b). 'Grammaticalization: Synchronic Variation and Diachronic Change, *Lingua e Stile* 20: 303–318.

Levinson, Stephen C. (2006). 'The Language of Space in Yélî Dnye', in Stephen C. Levinson and David P. Wilkins (eds.), *Grammars of Space: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press, 157–205.

Levy, Roger (2005). *Probabilistic Models of Word Order and Syntactic Discontinuity*. Ph.D. dissertation, Stanford University.

Levy, Roger (2008). 'Expectation-Based Syntactic Comprehension', *Cognition*, 106: 1126–1177.

Lewis, M. Paul (ed.) (2009). *Ethnologue: Languages of the World, Sixteenth Edition*. Dallas, TX: SIL International.

Lewis, M. Paul, Simons, Gary F., and Fennig, Charles D. (eds.) (2013). *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Lichtenberk, Frantisek (1983). *A Grammar of Manam*. Honolulu, HI: University of Hawaii Press.

Lidz, Jeffrey (2006). 'The Grammar of Accusative Case in Kannada', *Language* 82: 10–32.

Liégeois, Frédérique, Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., and Vargha-Khadem, F. (2003). 'Language fMRI Abnormalities Associated with FOXP2 Gene Mutation', *Nature Neuroscience* 6: 1230–1237.

Lillas-Schuil, Rosmari (2006). 'A Survey of Syntagms in the Hebrew Bible Classified as Hendiadys', in L. Edzard, and J. Retsö (eds.), *Current Issues in the Analysis of Semitic Grammar and Lexicon: Oslo-Göteborg Cooperation 4th-5th November 2005*. Göttingen: Otto Harrassowitz Verlag, 79–100.

Lind, Åge (1983). 'The Variant Forms of *try* and/*try to*', *English Studies* 5: 550–563.

Lindström, Eva (2002). *Topics in the Grammar of Kuot: A Non-Austronesian Language of New Ireland, Papua New Guinea*. Ph.D. dissertation, Stockholm University.

Lindström, Eva (2008). 'Language Complexity and Interlinguistic Difficulty', in Matti Miestamo, Kaius Sinnemäki and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam and Philadelphia: John Benjamins, 217–235.

LINGUIST List (2009). *Multitree: A Digital Library of Language Relationships*. Institute for Language Information and Technology (LINGUIST List), Eastern Michigan University. Ypsilanti, MI. http://multitree.org/.

Locke, John L. (2009). 'Evolutionary Developmental Linguistics: Naturalization of the Faculty of Language', *Language Sciences* 31: 33–59.

Lødrup, Helge (2002). 'The Syntactic Structures of Norwegian Pseudocoordinations', *Studia Linguistica* 56: 121–143.

Longobardi, Giuseppe (1994). 'Reference and Proper Names', *Linguistic Inquiry* 25: 609–665.

Lord, Carol (1989). *Syntactic Reanalysis in the Historical Development of Serial Verb Constructions in Languages of West Africa*. Ph.D. dissertation, University of California, Los Angeles.

Lötscher, Andreas (1978). 'Zur Verbstellung im Zürichdeutschen und in anderen Varianten des Deutschen', *Zeitschrift für Dialektologie und Linguistik* 45: 1–29.

Loughnane, Robyn (2009). *A Grammar of Oksapmin*. Ph.D. dissertation, University of Melbourne.

Lupyan, Gary, and Dale, Rick (2010). 'Language Structure is Partly Determined by Social Structure', *PLoS ONE* 5: 1–10.

Lynch, John (1983). 'Switch Reference in Lenakel', in J. Haiman, and P. Munro (eds.), *Switch Reference and Universal Grammar*. Amsterdam: John Benjamins, 209–221.

MacDonald, Maryellen C., Pearlmutter, Neal J., and Seidenberg, Mark S. (1994). 'The Lexical Nature of Syntactic Ambiguity Resolution', *Psychological Review* 101.4: 676–703.

MacDonald, Maryellen C., and Thornton, Robert (2009). 'When Language Comprehension Reflects Production Constraints: Resolving Ambiguities with the Help of Past Experience', *Memory and Cognition* 37: 1177–1186.

Maclean, Paul D. (1949). 'Psychosomatic Disease and the "Visceral Brain": Recent Developments Bearing on the Papez Theory of Emotion', *Psychosomatic Medicine* 11: 338–353.

Maddieson, Ian (1980). 'Phonological Generalizations from the UCLA Phonological Segment Inventory Database (UPSID)', *UCLA Working Papers in Phonetics* 50: 57–68.

Maddieson, Ian (1984). *Pattern of Sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian (2006). 'Correlating Phonological Complexity: Data and Validation', *Linguistic Typology* 10: 106–123.

Maddieson, Ian (2007). 'Issues of Phonological Complexity: Statistical Analysis of the Relationship Between Syllable Structures, Segment Inventories, and Tone Contrasts', in Maria-Josep Soleí, Patrice Speeter Beddor and Manjari Ohala (eds.), *Experimental Approaches to Phonology*. Oxford: Oxford University Press, 93–103.

Maddieson, Ian (2009). 'Calculating Phonological Complexity', in Iona Chitoran, Christophe Coupe and Egidio Marsico (eds.), *Approaches to Phonological Complexity*. Berlin: Mouton De Gruyter, 85–109.

Maddieson, Ian (2011a). 'Phonological Complexity in Linguistic Patterning', *ICPhS XVII*: 28–34.

Maddieson, Ian (2011b). 'Syllable Structure', in Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. http://wals.info/.

Maddieson, Ian, and Precoda, Kristin (1990). 'Updating UPSID', *UCLA Working Papers in Phonetics* 74: 104–111.

Mak, W. M., Vonk, W., and Schriefers, H. (2002). 'The Influence of Animacy on Relative Clause Processing', *Journal of Memory and Language* 47: 50–68.

Makuuchi, Michiru, Bahlmann, Jörg, Anwander, Alfred, and Friederici, Angela D. (2009). 'Segregating The Core Computational Faculty of Human Language from Working Memory', *Proceedings of The National Academy of Sciences* 106: 8362–8367.

Malchukov, Andrej, and Spencer, Andrew (2009). 'Typology of Case Systems', in Andrej Malchukov and Andrew Spencer (eds.), *The Oxford Handbook of Case*. Oxford: Oxford University Press, 651–668.

Manber, Udi (1989). *Introduction to Algorithms: A Creative Approach*. Reading, MA: Addison-Wesley.

Marcus, Gary, Pinker, Steven, Ullman, Michael, Hollander, Michelle, Rosen, John and Xu, Fei (1992). *Overregularization in Language Acquisition*. Chicago: University of Chicago Press.

Marková, Jana, and Cséfalvay, Szolt (2010). 'Morphological-syntactic Deficits in the Production of Slovak-speaking Aphasic Patients', *Aphasiology* 24: 1197–1222.

Marsaja, I Gede (2008). *Desa Kolok - A Deaf Village and its Sign Language in Bali, Indonesia*. Nijmegen: Ishara Press.

Martinet, André (1955). *Économie des changements phonétiques*. Bern: A. Francke.

Martinet, André (1962). A *Functional View of Language*. Oxford: Oxford University Press.

Maslova, Elena (2000). 'A Dynamic Approach to the Verification of Distributional Universals', *Linguistic Typology* 4: 307–333.

Matisoff, James A. (1973). 'Tonogenesis in Southeast Asia', in Larry M. Hyman (ed.), *Consonant Types and Tone*. Los Angeles: Southern California Occasional Papers in Linguistics 1, 71–95.

Matthewson, Lisa (1998). *Determiner Systems and Quantificational Strategies: Evidence from Salish.* The Hague: Holland Academic Graphics.

Matthewson, Lisa (1999). 'On the Interpretation of Wide-Scope Indefinites', *Natural Language Semantics* 7: 79–134.

Matthewson, Lisa (2001). 'Quantification and the Nature of Cross-Linguistic Variation', *Natural Language Semantics* 9: 145–189.

Matthewson, Lisa (2004). 'On the Methodology of Semantic Fieldwork', *International Journal of American Linguistics* 70: 369–415.

Matthewson, Lisa (2005). *When I Was Small - I Wan Kwikws: A Grammatical Analysis of St'át'imcets Oral Narratives.* Vancouver: UBC Press.

Matthewson, Lisa (2006a). 'Presuppositions and Cross-Linguistic Variation'. *Proceedings of the North East Linguistics Society* 36. Amherst, MA: GLSA.

Matthewson, Lisa (2006b). 'Temporal Semantics in a Supposedly Tenseless Language'. *Linguistics and Philosophy* 29: 673–713.

Matthewson, Lisa (2008). *Quantification: A Cross-Linguistic Perspective.* East Sussex: Emerald.

Matthewson, Lisa (2009). 'Pronouns, Presuppositions, and Semantic Variation'. *Proceedings of SALT XVIII*. Ithaca, NY: CLC Publications, 527–550.

Matthewson, Lisa (2011). 'Methods In Cross-Linguistic Semantics', in Claudia Maienborn, Klaus von Heusinger, and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, vol. 1. Berlin: De Gruyter, 268–285.

Matthewson, Lisa (2012). 'Evidence About Evidentials: Where Fieldwork Meets Theory', in Britta Stolterfoht and Sam Featherston (eds.), *Empirical Approaches To Linguistic Theory: Studies of Meaning and Structure*. Berlin: De Gruyter, 85–114.

Matthewson, Lisa (in press). 'Evidential Restrictions on Epistemic Modals', in Luis Alonso-Ovalle and Paula Menendez-Benito (eds.), *Epistemic Indefinites*. Oxford: Oxford University Press.

Matthewson, Lisa, Davis, Henry, and Rullmann, Hotze (2007). 'Evidentials as Epistemic Modals: Evidence From St'át'imcets'. *The Linguistic Variation Yearbook* 7: 201–254.

McCloy, Daniel, Moran, Steven, and Wright, Richard (2013). 'Revisiting "The role of Features in Phonological Inventories"', Paper presented at the CUNY Conference on the Feature in Phonology and Phonetics. New York.

McCrum, Robert, Cran, William, and MacNeil, Robert (1992). *The Story of English.* New York, NY: Penguin Books.

McGregor, William B. (1989). 'Discourse Basis of Ergative Marking in Gooniyandi', *La Trobe University Working Papers in Linguistics* 2: 127–158.

McGregor, William B. (1990). *A Functional Grammar of Gooniyandi*. Amsterdam: John Benjamins.

McKenzie, Andrew R. (2012). *The Role of Contextual Restriction in Reference Tracking*. Ph.D. dissertation, University of Massachusetts, Amherst.

McMahon, April M. S. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.

McWhorter, John H. (2001). 'The World's Simplest Grammars Are Creole Grammars', *Linguistic Typology* 5: 125–166.

McWhorter, John H. (2005). *Defining Creole*. Oxford and New York: Oxford University Press.

McWhorter, John H. (2007). *Language Interrupted: Signs of Non-native Acquisition in Standard Language Grammars*. New York: Oxford University Press.

McWhorter, John H. (2011). *Linguistic Simplicity and Complexity*: *Why Do Languages Undress?* Berlin: De Gruyter.

Meier, Richard P. (1984). 'Sign as Creole', *Behavioral and Brain Sciences* 7: 201–202.

Meier, Richard P. (2002). 'Why Different, Why the Same? Explaining Effects and Non-Effects of Modality upon Linguistic Structure in Sign and Speech', in R. P. Meier, K. Cormier and D. Quinto-Pozos (eds.), *Modality and Structure in Signed and Spoken Languages*. Cambridge: Cambridge University Press, 1–25.

Menges, Karl. H. (1975). *Altajische Studien II. Japanisch und Altajisch*. Wiesbaden: Franz Steiner Verlag.

Menn, Lise, and Duffield, Cecily Jill (2013). 'Aphasias and Theories of Linguistic Representation: Representing Frequency, Hierarchy, Constructions, and Sequential Structure', *WIREs: Cognitive Science* 4:651–663. doi: 10.1002/wsc.1257.

Menn, Lise, and Obler, Loraine K. (1990). 'Chapter 20, Conclusion: Cross-language Data and Theories of Agrammatism', in L. Menn and L. K. Obler (eds.), *Agrammatic Aphasia*, vol. 2, 1369–1389.

Michaelis, Susanna, Maurer, Philippe, Haspelmath, Martin and Huber, Magnus (eds.) (2013). *The Atlas of Pidgin and Creole Language Structures* (APiCS), Oxford: Oxford University Press.

Michel, Jean-Baptiste, Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., and Aiden, E.L. (2011). 'Quantitative Analysis of Culture Using Millions of Digitized Books', *Science* 331: 176–182. Google Ngrams corpus available at http://books.google.com/ngrams/.

Miestamo, Matti (2005). *Standard Negation: The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin: Mouton de Gruyter.

Miestamo, Matti (2006a). 'On the Feasibility of Complexity Metrics', in K. Kerge and M.-M. Sepper (eds.), *FinEst Linguistics, Proceedings of the Annual Finnish and Estonian Conference of Linguistics*. Tallinn, Estonia, 11–26.

Miestamo, Matti (2006b). 'On the Complexity of Standard Negation', in Mickael Suominem, Antti Arppe, Anu Airola, Orvokki Heinämäki, Matti Miestamo, Urho Määttä, Jussi Niemi, Kari K. Pitkänen, and Kaius Sinnemäki (eds.), *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday*. Turku: The Linguistic Association of Finland, 345–356.

Miestamo, Matti (2008). 'Grammatical Complexity in a Cross-Linguistic Perspective', in Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 23–41.

Miestamo, Matti (2009). 'Implicational Hierarchies and Grammatical Complexity', in Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 80–97.

Miestamo, Matti, Sinnemäki, Kaius, and Karlsson, Fred (eds.) (2008). *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins.

Miles, H. Lyn White (1990). 'The Cognitive Foundations for Reference in a Signing Orangutan', in S. T. Parker and K. R. Gibson (eds.), *'Language' and Intelligence in Monkeys and Apes: Comparative Developmental Perspectives*. Cambridge: Cambridge University Press, 511–539.

Miller, George A. (1956). 'The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information', *The Psychological Review* 63: 81–97.

Miller, George A. and Chomsky, Noam (1963). 'Finitary Models of Language Users', in P. Luce, R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, volume 2. New York: Wiley, 419–492.

Milsark, Gary (1974). *Existential Sentences in English*. Ph.D. dissertation, MIT.

Mistry, P. J. (1998). Objecthood and Specificity in Gujarati, in J.H. Hill, P.J. Mistry and L. Campbell (eds.), *The Life of Language, Papers in Honor of William Bright*. Berlin: Mouton de Gruyter, 425–442.

Mithun, Marianne (1991). 'Active/Agentive Case Marking and its Motivations', *Language* 67: 510–546.

Miyagawa, Shigeru (2010). *Why Agree? Why Move? Unifying Agreement-Based and Discourse-Configurational Languages.* Cambridge, MA: MIT Press.

Moag, Rodney (1977). *Fiji Hindi: A Basic Course and Reference Grammar*. Canberra: Australian National University.

Mobbs, Iain (2008). '"Functionalism", the Design of the Language Faculty, and (Disharmonic) Typology', MS, University of Cambridge. http://ling.auf.net/lingBuzz/000680 (Accessed 5th March 2014).

Mohammed, Mohammed A. (2001). *Modern Swahili Grammar*. Nairobi, Kenya: East African Educational Publishers Ltd.

Montague, Richard (1974). 'The Proper Treatment of Quantification in Ordinary English', in Thomason, Richmond (ed.), *Formal Philosophy*. New Haven: Yale University Press, 247–270.

Moran, Steven (2012). *Phonetics Information Base and Lexicon*. Ph.D. dissertation, University of Washington.

Moran, Steven, McCloy, Daniel, and Wright, Richard (2012). 'Revisiting Population Size vs. Phoneme Inventory Size', *Language* 88: 877–893.

Moravcsik, Edith (2011). 'Explaining Language Universals', in Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford: Oxford University Press, 69–89.

Moro, Andrea (2008). *The Boundaries of Babel: The Brain and the Enigma of Impossible Languages*. Cambridge, MA: MIT Press.

Moro, Andrea, Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., and Fazio, F. (2001). 'Syntax and the Brain: Disentangling Grammar by Selective Anomalies', *Neuroimage* 13: 110–118.

Müller-Gotama, Frans (1994). *Grammatical Relations: A Cross-Linguistic Perspective on their Syntax and Semantics*. Berlin: De Gruyter.

Munn, Alan Boag (1993). *Topics in the Syntax and Semantics of Coordinate Structures*. Ph.D. dissertation, University of Maryland.

Murane, Elizabeth (1974). *Daga Grammar, from Morpheme to Discourse*. Norman, OK: Summer Institute of Linguistics of the University of Oklahoma.

Muskens, Reinhard (2007). 'Separating Syntax and Combinatorics in Categorial Grammar', *Research on Language and Computation* 5: 267–285.

Musso, Mariacristina, Moro, Andrea, Glauche, Volkmar, Rijntjes, Michel, Reichenbach, Jürgen, Büchel, Christian and Weiller, Cornelius (2003). 'Broca's Area and the Language Instinct', *Nature Neuroscience* 6: 774–781.

Muysken, Pieter (1988). 'Are Creoles a Special Type of Language?', in Frederick J. Newmeyer (ed.), *Linguistics: The Cambridge Survey, volume 2: Linguistic Theory: Extensions and Implications*. Cambridge: Cambridge University Press, 285–301.

Nash, Lea (1995). *Argument Scope and Case Marking in SOV and in Ergative Languages: The Case of Georgian*. Ph.D. dissertation, Université Paris 8.

Nash, Lea (1996). 'The Internal Ergative Subject Hypothesis', *Proceedings of NELS 26*: 195–210.

Neeleman, Ad, and Szendrői, Kriszta (2004). 'Superman Sentences', *Linguistic Inquiry* 35: 149–159.

Neeleman, Ad, and Szendrői, Kriszta (2007). 'Radical Pro-Drop and the Morphology of Pronouns', *Linguistic Inquiry* 38: 671–714.

Nettle, Daniel (1995). 'Segmental Inventory Size, Word Length, and Communicative Efficiency', *Linguistics* 33: 359–367.

Nevins, Andrew, Pesetsky, David, and Rodrigues, Cilene (2009). 'Pirahã Exceptionality: A Reassessment', *Language* 85: 355–404.

Newmeyer, Frederick J. (1991). 'Functional Explanation in Linguistics and the Origin of Language', *Language and Communication* 11: 1–28.

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. (2004). 'Against a Parameter-Setting Approach to Language Variation', *Linguistic Variation Yearbook* 4: 181–234.

Newmeyer, Frederick J. (2005). *Possible and Probable Languages. A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Newmeyer, Frederick J. (2007). '"More Complicated and Hence, Rarer": A Look at Grammatical Complexity and Cross-linguistic Rarity', in S. Karimi, V. Samiian, and W. K. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*. Amsterdam and Philadelphia: Benjamins, 221–242.

Newmeyer, Frederick J. (2009). 'On Split CPs and the "Perfectness" of Language', in B. Shaer, P. Cook, W. Frey, and C. Maienborn (eds.), *Dislocated Elements in Discourse: Syntactic, Semantic, and Pragmatic Perspectives*. New York and London: Routledge, 114–140.

Newmeyer, Frederick J. (2011). 'Can One Language be "More Complex" than Another?', unpublished MS, University of Washington, University of British Columbia, and Simon Fraser University.

Newport, Elissa L., and Suppala, Ted (2000). 'Sign Language Research at the Millennium', in K. Emmorey and H. Lane (eds.), *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*. Mahway NJ: Lawrence Erlbaum Associates, 103–114.

Nichols, Johanna (1992). *Linguistic Diversity in Space and Time.* Chicago: University of Chicago Press.

Nichols, Johanna (2009). 'Linguistic Complexity: A Comprehensive Definition and Survey', in Geoffrey Sampson, David Gil and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 64–79.

Nichols, Johanna, and Bickel, Balthasar (2005). 'Locus of Marking in the Clause', in M. Haspelmath, M. Dryer, D. Gil and B. Comrie (eds.), *The World Atlas of Language Structures*, Oxford: Oxford University Press, 98–101.

Nicol, Janet L., Jakubowicz, Célia, and Goldblum, Marie-Claire (1996). 'Sensitivity to Grammatical Marking in English-speaking and French-speaking non-fluent Aphasics', *Aphasiology* 10: 593–622.

Nonaka, Angela M. (2007). *Emergence of an Indigenous Sign Language and a Speech/Sign Community in Ban Khor*, Thailand, Ph.D. dissertation, UCLA.

Nowak, Andrzej, and Latané, Bibb (1994). 'Simulating the Emergence of Social Order From Individual Behaviour', in N. Gilbert and J. Doran (eds.), *Simulating Societies: The Computer Simulation of Social Phenomena.* London: UCL Press.

Nyombe, Bureŋ G.V. (1987). *Argument Bearing Affixes in Bari*, Ph.D. dissertation, City University of New York.

Nyst, Victoria (2007). *A Descriptive Analysis of Adamorobe Sign Language*. Utrecht: LOT.

Oldfield, R. C., and Wingfield, Arthur (1965). 'Response Latencies in Naming Objects', *Quarterly Journal of Experimental Psychology* 17: 273–281.

Ordóñez, Francisco (2000). *The Clausal Structure of Spanish*. New York: Garland.

Ouali, Hamid (2006). *Unifying Agreement Relations: A Minimalist Analysis of Berber*. Ph.D. dissertation, University of Michigan.

Ouhalla, Jamal (1991). *Functional Categories and Parametric Variation*. London: Routledge and Kegan Paul.

Oxford English Dictionary Online (OED) (2013). Retrieved from Oxford English Dictionary Online: http://www.oed.com.

Palmer, Frank R. (1994). *Grammatical Roles and Relations*. Cambridge: Cambridge University Press.

Parker, Steve (2009). 'The Transparency and Density of Phonological Rule Interactions', *Linguistic Typology* 13: 197–265.

Parkvall, Mikael (2008). 'The Simplicity of Creoles in a Typological Perspective', in F. Karlsson, M. Miestamo and K. Sinnemäki (eds.), *Language Complexity: Typology, Contact, Change*, John Benjamins, Amsterdam, 265–285.

Partee, Barbara (1988). 'Many Quantifiers', in J. Powers and K. de Jong (eds.), *Proceedings of the Fifth ESCOL.* Columbus: The Ohio State University.

Partee, Barbara (1995). 'Quantificational Structures and Compositionality', in Bach et al. (eds.), *Quantification in Natural Languages*. Dordrecht: Kluwer, 541–601.

Partee, Barbara, ter Meulen, Alice, and Wall, Robert E. (1990). *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.

Patocka, Franz (1997). *Satzgliedstellung in Den Bairischen Dialekten Österreichs*, Bern: Peter Lang.

Payne, Doris L. (1985). *Aspects of the Grammar of Yagua: A Typological Perspective*. Ph.D. dissertation, University of California, Los Angeles.

Payne, Doris L., and Payne, Thomas E. (1990). 'Yagua', in Desmond C. Derbyshire and Geoffrey K. Pullum (eds.), *Handbook of Amazonian Languages*, vol. 2. Berlin: Mouton de Gruyter, 249–474.

Pellegrino, François, Coupé, Christophe, and Marsico, Egidio (2011). 'A Cross-language Perspective on Speech Information Rate', *Language* 87: 539–558.

Penchoen, Thomas G. (1973). *Tamazight of the Ayt Ndhir*. Los Angeles, CA: Undena Publications.

Peng, Anne (2011). 'Head-Final and Head-Initial Relative Clauses in Jambi Teochew', in K. Otaki, H. Takeyasu and S. Tanigawa (eds.), *Online Proceedings of GLOW in Asia Workshop for Young Scholars 2011.* http://faculty.human.mie-u.ac.jp/~glow_mie/Workshop_Proceedings/20Peng.pdf (Accessed 5th March 2014).

Perkins, Revere D. (1992). *Deixis, Grammar, and Culture*. Amsterdam: Benjamins.

Perlmutter, David (1978). 'Impersonal Passive and the Unaccusative Hypothesis', *Berkeley Linguistics Society* 4: 159–189.

Phimsawat, On-Usa (2011). *The Syntax of Pro-Drop in Thai*. Ph.D. dissertation, Newcastle University.

Picallo, M. Carme (1991). 'Nominals and Nominalizations in Catalan'. *Probus* 3: 279–316.

Pinker, Steven (1984). *Language Learnability and Language Development.* Cambridge, MA: Harvard University Press.

Pinker, Steven, and Bloom, Paul (1990). 'Natural Language and Natural Selection', *Behavioral and Brain Sciences* 13: 707–784.

Platzak, Christer (1990). 'A Grammar Without Functional Categories: A Syntactic Study of Early Child Language', *Nordic Journal of Linguistics* 13: 107–126.

Poeppel, David, and Embick, David (2005). 'Defining the Relation Between Linguistics and Neuroscience,' in A. Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones.* Mahwah, NJ: Lawrence Erlbaum, 103–118.

Pollock, Jean-Yves. (1989). 'Verb Movement, Universal Grammar and the Structure of IP', *Linguistic Inquiry* 20: 365–424.

Potts, Christopher (2005). *The Logic of Conventional Implicatures.* Oxford: Oxford University Press.

Potts, Christopher, and Roeper, Tom (2006). 'The Narrowing Acquisition Path: From Expressive Small Clauses to Declaratives', in L. Progovac, K. Paesani, E. Casielles, and E. Barton (eds.), *The Syntax of Nonsententials: Multidisciplinary Perspectives*. Amsterdam: Benjamins, 183–201.

Poutsma, Hendrik (1917). 'Hendiadys in English: Together with Some Observations on the Construction of Certain Verbs I', *Neophilologus* 2: 202–218.

Primus, Beatrice (1999). *Cases and Thematic Roles: Ergative, Accusative and Active*. Tübingen: Max Niemeyer Verlag.

Progovac, Ljiljana (2005). *A Syntax of Serbian: Clausal Architecture*. Bloomington, IN: Slavica Publishers.

Progovac, Ljiljana (2006). 'The Syntax of Nonsententials: Small Clauses and Phrases at the Root', in L. Progovac, K. Paesani, E. Casielles, and E. Barton (eds.), *The Syntax of Nonsententials: Multidisciplinary Perspectives*. Amsterdam: John Benjamins, 33–71.

Progovac, Ljiljana (2008a). 'Root Small Clauses with Unaccusative Verbs', in A. Antonenko, J. F. Bailyn, and C. Y. Bethin (eds.), *Proceedings of FASL (Formal Approaches to Slavic Linguistics)* 16. Ann Arbor: Michigan Slavic Publications, 359–373.

Progovac, Ljiljana (2008b). 'What Use Is Half a Clause?' In A. D. M. Smith, *et al.* (eds.), *Evolution of Language: Proceedings of EVOLANG7, Barcelona, Spain*. New Jersey: World Scientific, 259–266.

Progovac, Ljiljana (2009). 'Layering of Grammar: Vestiges of Evolutionary Development of Syntax in Present-Day Languages', in G. Sampson, D. Gil, and P. Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 203–212.

Progovac, Ljiljana (2010a). 'When Clauses Refuse to be Recursive: An Evolutionary Perspective', in H. Van Der Hulst (ed.), *Recursion in Human Language*. The Hague: Mouton De Gruyter, 193–211.

Progovac, Ljiljana (2010b). 'Syntax: Its Evolution and its Representation in the Brain', *Biolinguistics* 4: 233–254.

Progovac, Ljiljana (2012a). 'Compounds and Commands in the Evolution of Human Language', *Theoria et Historia Scientiarum: An International Journal for Interdisciplinary Studies* IX: 49–70.

Progovac, Ljiljana (2012b). 'Rigid Syntax, Rigid Sense'. Paper presented at FASL (Formal Approaches to Slavic Linguistics), Indiana University, Bloomington, May 2012.

Progovac, Ljiljana, and Locke, John L. (2009). 'The Urge to Merge: Ritual Insult and the Evolution of Syntax', *Biolinguistics* 3: 337–354.

Pullum, Geoffrey K. (1990). 'Constraints on Intransitive Quasi-serial Verb Constructions in Modern Colloquial English', *Ohio State University Working Papers in Linguistics* 39: 218–246.

Pullum, Geoffrey K., and Gazdar, Gerald (1982). 'Natural Languages and Context-Free Languages', *Linguistics and Philosophy* 4: 471–504.

Pullum, Geoffrey K., and Zwicky, Arnold M. (1986). 'Phonological Resolution of Syntactic Feature Conflict', *Language,* 62: 751–773.

Quer, Josep, and GRIN ('Gramàtica i interpretació de llengua de signes', Federació de Persones Sordes de Catalunya (FESOCA) and Universitat de Barcelona) (2008). 'Structures of Possession and Existence in Catalan Sign Language', in U. Zeshan and P. Perniss (eds.) *Possessive and Existential Constructions in Sign Languages*. Nijmegen: Ishara, 33–53.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. http://www.R-project.org (5th March 2014).

Radford, Andrew (1988). 'Small Children's Small Clauses', *Transactions of the Philological Society* 86: 1–43.

Radford, Andrew (1990). *Syntactic Theory and the Acquisition of English Syntax*. Oxford: Blackwell.

Radford, Andrew (2004). *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.

Ramchand, Gillian, and Svenonius, Peter (2013). 'Deriving the Functional Hierarchy'. Paper presented at GLOW 36, Lund University, April 2013.

Rasolofo, Andoveloniaina (2006). *Malagasy Transitive Clause Types and their Functions*. Ph. D. dissertation, University of Oregon.

Reich, Peter A. (1969). 'The Finiteness of Natural Language', *Language* 45: 831–843.

Rescher, Nicholas (1998). *Complexity: A Philosophical Overview*. New Brunswick: Transaction.

Rice, Keren (1989). *A Grammar of Slave*. Berlin: Mouton de Gruyter.

Richards, Marc (2004). *Object Shift and Scrambling in North and West Germanic. A Case Study in Symmetrical Syntax*. Ph.D. dissertation, University of Cambridge.

Riddle, Elizabeth M. (2008). 'Complexity in Isolating Languages: Lexical Elaboration vs. Grammatical Economy', in Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 133–151.

Ridley, Mark (1993). *Evolution*. Oxford: Blackwell Scientific Publications.

Ridley, Matt (1999). *Genome: The Autobiography of a Species in 23 Chapters*. London: Fourth Estate.

Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi (1994). 'Some Remarks on Linguistic Theory and Language Development: The Case of Root Infinitives', *Language Acquisition* 4: 371–393.

Rizzi, Luigi (1997). 'The Fine Structure of the Left Periphery', in L. Haegeman (ed.), *Elements of Grammar: Handbook in Generative Syntax*. Dordrecht: Kluwer, 281–337.

Rizzi, Luigi (2001). 'Relativized Minimality Effects', in M. Baltin and C. Collins (eds.), *A Handbook of Syntactic Theory*. Oxford: Blackwell, 89–110.

Rizzi, Luigi (2004). 'On the Cartography of Syntactic Structures', in L. Rizzi (ed.), *The Structure of CP and IP*. Oxford: Oxford University Press, 3–15.

Rizzi, Luigi (2013a). 'Locality', *Lingua* 130: 169–186.

Rizzi, Luigi (2013b). 'Notes on Cartography and Further Explanation', *Probus* 25: 197–226.

Roberts, Ian (1999). 'Verb Movement and Markedness', in Michel DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, MA: MIT Press, 287–327.

Roberts, Ian (2007). *Diachronic Syntax*. Oxford: Oxford University Press.

Roberts, Ian (2010). *Agreement and Head Movement. Clitics and Defective Goals*. Cambridge, MA: MIT Press.

Roberts, Ian (2012). 'Macroparameters and Minimalism: A Programme for Comparative Research', in C. Galves, S. Cyrino, R. Lopes, F. Sândalo and J. Avelar (eds.), *Parameter Theory and Linguistic Change*. Oxford: Oxford University Press, 319–334.

Roberts, Ian (to appear). 'Some Comments on "The Structure of Syntactic Typologies" ', *Mind and Language*.

Roberts, Ian and Holmberg, Anders (2010). 'Introduction: Parameters in Minimalist Theory', in T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press, 1–57.

Roberts, Ian, and Roussou, Anna (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.

Rohdenburg, Günter (1974). *Sekundäre Subjektivierungen im Englischen und Deutschen*. Bielefeld: Cornelson-Velhagen und Klasing.

Rohdenburg, Günter (1996). 'Cognitive Complexity and Grammatical Explicitness in English', *Cognitive Linguistics* 7: 149–182.

Rolfe, Leonard (1996). 'Theoretical Stages in the Prehistory of Grammar', in A. Lock and C. R. Peters (eds.), *Handbook of Human Symbolic Evolution*. Oxford: Clarendon Press, 776–792.

Ross, Daniel (2013a). 'Verbal Pseudocoordination in English: A Syntactic Analysis with Reference to Diachronic, Dialectal, and Crosslinguistic Variation'. Unpublished qualifying exam paper, Linguistics Department, University of Illinois at Urbana-Champaign. https://www.ideals.illinois.edu/handle/2142/42581 (Accessed 5th March 2014).

Ross, Daniel (2013b). 'Dialectal Variation and Diachronic Development of *try*-complementation'. *Studies in the Linguistic Sciences: Illinois Working Papers: 108–47.* http://hdl.handle.net/2142/46461.

Ross, John R. (1967). *Constraints on Variables in Syntax*, Ph.D. dissertation, MIT.

Rounds, Carol (2001). *Hungarian: An Essential Grammar*. London: Routledge.

Ruhlen, Merritt (1975). *A Guide to the Languages of the World*. Stanford University: Language Universals Project.

Rullmann, Hotze, Matthewson, Lisa, and Davis, Henry (2008). 'Modals as Distributive Indefinites', *Natural Language Semantics* 16: 317–357.

Sachs, Jacqueline (1967). 'Recognition Memory for Syntactic and Semantic Aspects of Connected Discourse', *Perception and Psychophysics* 2: 437–442.

Saeed, John Ibrahim (1999). *Somali*. Amsterdam: John Benjamin.

Saffran, Jenny R. (2003). 'Statistical Language Learning: Mechanisms and Constraints', *Current Directions in Psychological Science,* 12: 110–114.

Sag, Ivan A. (1997). 'English Relative Clause Constructions', *Journal of Linguistics* 33: 431–484.

Sag, Ivan A. (2012). 'Sign-Based Construction Grammar: An Informal Synopsis', in Hans C. Boas and Ivan A. Sag (eds.), *Sign-Based Construction Grammar*. Stanford: CSLI Publications. [Distributed by U. of Chicago Press].

Sag, Ivan A., and Wasow, Thomas (2011). 'Performance-Compatible Competence Grammar', in Robert Borsley and Kersti Borjars (eds.), *Non-Transformational Syntax: Formal and Explicit Models of Grammar*. Oxford: Wiley-Blackwell.

Sag, Ivan A., Gazdar, Gerald, Wasow, Thomas, and Weisler, Steven (1985). 'Coordination and How to Distinguish Categories', *Natural Language and Linguistic Theory* 3: 117–171.

Saito, Mamoru (2007). 'Notes on East Asian Argument Ellipsis', *Language Research* 43: 203–227.

Sampson, Geoffrey (2009a). 'An Interview with Dan Everett', in Geoffrey Sampson, David Gil, and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 213–229.

Sampson, Geoffrey (2009b). 'A Linguistic Axiom Challenged', in Geoffrey Sampson, David Gil and Peter Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 1–18.

Sampson, Geoffrey, Gil, David, and Trudgill, Peter (eds.) (2009). *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.

Sandler, Wendy, and Lillo-Martin, Diane (2005). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Sandler, Wendy, Meir, Irit, Padden, Carol, and Aronoff, Mark (2005). 'The Emergence of Grammar in a New Sign Language', *Proceedings of the National Academy of Sciences* 102, No. 7: 2661–2665.

Santi, Andrea, and Grodzinsky, Yosef (2007). 'Working Memory and Syntax Interact in Broca's Area', *Neuroimage* 37: 8–17.

Santi, Andrea, and Grodzinsky, Yosef (2010). 'fMRI Adaptation Dissociates Syntactic Complexity Dimensions', *Neuroimage* 51: 1285–1293.

Sapir, Edward (1912). 'Language and Environment', *American Anthropologist* 14: 226–242.

Sapir, Edward (1921). *Language*. New York: Harcourt, Brace, and World.

Sapp, Christopher D. (2005). 'Factors Favoring Aux-V Order in 15th-Century German Subordinate Clauses', in Dorian Roehrs, Ock-Hwan Kim, and Yoshihisa Kitagawa (eds.), *Syntax and Beyond: Indiana University Working Papers in Linguistics*, vol. 5. Bloomington: Indiana University Linguistics Club Publications.

Sapp, Christopher D. (2006). *Verb Order in Subordinate Clauses From Early New High German to Modern German*. Ph.D. dissertation, Indiana University.

Sapp, Christopher D. (2007). 'Focus and Verb Order in Early New High German: Historical and Contemporary Evidence', in Sam Featherston and Wolfgang Sternefeld (eds.), *Roots: Linguistics in Search of Its Evidential Base*. Berlin: Walter de Gruyter.

Sapp, Christopher D. (2011). *The Verbal Complex in Subordinate Clauses from Medieval to Modern German*. Amsterdam: John Benjamins.

Sasanuma, Sumiko, Kamio, Akio, and Kubota, Masahito (1990). 'Agrammatism in Japanese: Two Case Studies', in Lise Menn and Loraine K. Obler (eds.), *Agrammatic Aphasia: A Cross-language Narrative Sourcebook*. Amsterdam: John Benjamins, 1225–1283.

Sauerland, Uli (2014). 'Against Complexity Parameters', in A. Trotzke and J. Bayer (eds.), *Syntactic Complexity Across Interfaces*. Berlin: Mouton de Gruyter.

Sauerland, Uli, and Trotzke, Andreas (2011). 'Biolinguistic Perspectives on Recursion: Introduction to the Special Issue', *Biolinguistics* 5: 1–9.

Scha, Remko (1981). 'Distributive, Collective and Cumulative Quantification', in J.A. G. Groenendijk, T.M.V. Janssen and M.B.J. Stokhof (eds.), *Formal Methods in the Study of Language* (Mathematical Centre Tracts 136). Amsterdam: University of Amsterdam, 483–512.

Schirrmeister, Bettina E., Antonelli, Alexandre, and Bagheri, Homayoun C. (2011). 'The Origin of Multicellularity in Cyanobacteria', *BMC Evolutionary Biology* 11.45: 1–21.

Schlenker, Philippe (2005). 'Non-Redundancy: Towards a Semantic Reinterpretation of Binding Theory', *Natural Language Semantics* 13: 1–92.

Schmid, Tanja (2005). *Infinitival Syntax. Infinitivus Pro Participio as a Repair Strategy*. Amsterdam: John Benjamins.

Schmid, Tanja, and Vogel, Ralf (2004). 'Dialectal Variation in German 3-Verb Clusters: A Surface-Oriented Optimality Theoretic Account', *The Journal of Comparative Germanic Linguistics* 7: 235–274.

Schreuder, Robert, Grende, Marjon Poulisse, Nanda Roelofs, Ardi, and van de Voort, Marlies (1990). 'Lexical Processing, Morphological Complexity, and Reading', *Comprehension Processes in Reading*. Hillsdale, NJ: Erlbaum, 125–141.

Schuchardt, Hugo (1980). *Pidgin and Creole Languages: Selected Essays (Edited and Translated by G. G. Gilbert)*. Cambridge: Cambridge University Press.

Schütz, Albert J. (1985). *The Fijian Language*, Honolulu: University of Hawaii Press.

Sedivy, Julie (2007). 'Implicature During Real Time Conversation: A View From Language Processing Research', *Philosophy Compass* 2/3: 475–496.

Seidenberg, Mark, and Petitto, Laura (1978). 'Signing Behavior In Apes: A Critical Review', *Cognition* 17: 177–215.

Seiler, Guido (2004). 'On Three Types of Dialect Variation and their Implications for Linguistic Theory: Evidence From Verb Clusters in Swiss German Dialects', in Bernd Kortmann (ed.), *Dialectology Meets Typology: Dialect Grammar From a Cross-Linguistic Perspective.* Berlin: Mouton de Gruyter.

Seiler, Guido (2008). 'Syntaxgeographie und Plastizität der Grammatik', in Karin Donhauser, Elvira Glaser, and Marcel Vuillaume (eds.), *Empirische Grundlagen Moderner Grammatikforschung: Akten Des XI. Internationalen Germnistenkongresses Paris 2005, Vol 4. Jahrbuch für Internationale Germanistik 80,* Bern: Peter Lang.

Selkirk, Elizabeth (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elizabeth (2000). 'The Interaction of Constraints on Prosodic Phrasing', in Merle Horne (ed.), *Prosody: Theory and Experiment*. Dordrecht: Kluwer, 231–261.

Senghas, Ann (2003). 'Intergenerational Influence and Ontogenetic Development in the Emergence of Spatial Grammar in Nicaraguan Sign Language', *Cognitive Development* 18: 511–531.

Shalizi, Cosma Rohilla (2006). 'Methods and Techniques of Complex Systems Science: An Overview', in T. Deisboeck and J. Kresh (eds.), *Complex Systems Science in Biomedicine*. New York: Springer, 33–114.

Shannon, Claude E. (1948). 'A Mathematical Theory of Communication', *Bell System Technical Journal* 27: 379–423, 623–656.

Sharpe, Margaret (1972). *Alawa Phonology and Grammar*. Canberra: Australian Institute of Aboriginal Studies.

Shearin, Hubert Gibson (1903). *The Expression of Purpose in Old English Prose.* New York: Henry Holt.

Sheehan, Michelle (2013a). 'Explaining the Final-Over-Final Constraint: Formal and Functional Approaches', in T. Biberauer and M. Sheehan (eds.). *Theoretical Approaches To Disharmonic Word Orders*. Oxford: Oxford University Press.

Sheehan, Michelle (2013b). 'Some Implications of a Copy Theory of Labeling', *Syntax* 16(4): 362–396.

Sheehan, Michelle (2014). 'Towards a General Alignment Parameter Hierarchy', to appear in N. Danton, D. Kostadinovska and R. Santana-Labarge (eds.), *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL 31).* Somerville, MA: Cascadilla.

Shetreet, Einat, Friedmann, Na'ama, and Hadar, Uri (2009a). 'An fMRI Study of Syntactic Layers: Sentential and Lexical Aspects of Embedding', *Neuroimage* 48: 707–716.

Shetreet, Einat, Friedmann, Na'ama, and Hadar, Uri (2009b). 'The Neural Correlates of Linguistic Distinctions: Unaccusative and Unergative Verbs', *Journal of Cognitive Neuroscience* 22: 2306–2315.

Shieber, Stuart M. (1987). 'Evidence Against the Context-Freeness of Natural Language', in Walter J. Savitch, Emmon Bach, William Marsh, and Gila Safran-Naveh (eds.) *The Formal Complexity of Natural Language*. Dordrecht: Reidel, 320–334.

Shosted, Ryan K. (2006). 'Correlating Complexity: A Typological Approach', *Linguistic Typology* 10: 1–40.

Siewierska, Anna (1998). 'Variation in Major Constituent Order: A Global and a European Perspective', in Anna Siewierska (ed.), *Constituent Order in The Languages of Europe: Empirical Approaches to Language Typology*. Berlin: Mouton De Gruyter, 475–551.

Sigurðsson, Halldor (2011). 'Uniformity and Diversity: A Minimalist Perspective', *Linguistic Variation Yearbook* 11: 189–222.

Simon, Herbert A. (1996). *The Sciences of the Artificial* (3rd edn.). Cambridge, MA: MIT Press.

Simpson, Andrew (2005). 'Classifiers and DP Structure in Southeast Asia', in R. Kayne and G. Cinque (eds.), *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press, 806–38.

Sinnemäki, Kaius (2008). 'Complexity Trade-Offs in Core Argument Marking', in Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 67–88.

Sinnemäki, Kaius (2011). *Language Universals and Linguistic Complexity: Three Case Studies in Core Argument Marking*. Ph.D. dissertation, University of Helsinki.

Sippola, Eva (2011). *Una gramática descriptiva del chabacano de Ternate*, Ph.D. dissertation, University of Helsinki.

Smith, Ian (1979). 'Substrata vs. Universals in the Formation of Sri Lanka Portuguese', in P. Mühlhäusler *et al.* (eds.), *Papers in Pidgin and Creole Linguistics No. 2*, Pacific Linguistics Series A - No. 57, Australian National University, Canberra, 183–200.

Smith, Neil (1999). *Chomsky: Ideas and Ideals*. Cambridge: Cambridge University Press.

Smith, Wayne H. (1990). 'Evidence for Auxiliaries in Taiwan Sign Language', in S.D. Fischer and P. Siple (eds.), *Theoretical Issues in Sign Language Research 1*. Chicago: University of Chicago Press, 211–228.

Smolensky, Paul, and Legendre, Géraldine (2006). 'Harmony Optimization and the Computational Architecture of the Mind/Brain', in *The Harmonic Mind, Vol.1: Cognitive Architecture*. Cambridge MA: MIT Press, 3–56.

Soh, Hooi Ling (1998). *Object Scrambling in Chinese*. Ph.D. dissertation, MIT.

Speas, Margaret (1990). *Phrase Structure in Natural Language*. Dordrecht: Kluwer.

Spivey, Michael J., Tanenhaus, Michael K., Eberhard, Kathleen M., and Sedivy, Julie C. (2002). 'Eye Movements and Language Comprehension: Effects of Visual Context on Syntactic Ambiguity Resolution', *Cognitive Psychology* 45: 447–481.

Sridhar, Shikaripur N. (1990). *Kannada*. London: Routledge.

Stalnaker, Robert (1974). 'Pragmatic Presuppositions', in Milton Munitz and Peter Unger (eds.), *Semantics and Philosophy*. New York: New York University Press, 197–213.

Starke, Michal (2001). *Move Reduces To Merge: A Theory of Locality*. Ph.D. dissertation, University of Geneva.

Staudacher-Valliamée, Gillette (2004). *Grammaire du créole réunionnais*. Paris: Sedes, Université de La Réunion.

Steever, Stanford B. (2009). 'Tamil and the Dravidian Languages', in B. Comrie (ed.), *The World's Major Languages* (2nd edition). London: Routledge, 635–652.

Stowell, Tim (1981). *Origins of Phrase Structure*. Ph.D. dissertation, MIT.

Stowell, Tim (1983). 'Subjects Across Categories', *The Linguistic Review* 2/3: 285–312.

Stringer, Christopher B., and Andrews, Peter (1988). 'Genetic and Fossil Evidence for the Origin of Modern Humans', *Science* 239: 1263–1268.

Stromswold, Karin, Caplan, David, Alpert, Nathaniel, and Rauch, Scott (1996). 'Localization of Syntactic Comprehension by Positron Emission Tomography', *Brain and Language* 52: 452–473.

Studdert-Kennedy, Michael (1991). 'Language Development from an Evolutionary Perspective', in N. A. Krasnegor, D. M. Rumbaugh, R. L. Scheiefelbusch, and M. Studdert-Kennedy (eds.), *Biological and Behavioral Determinants of Language Development*. Hillsdale, NJ: Erlbaum, 5–28.

Szabolcsi, Anna (2010). *Quantification*. Cambridge: Cambridge University Press.

Tanner, Darren (2012). 'Structural Effects in Agreement Processing: ERP and Reaction Time Evidence for Comprehension/Production Asymmetries'. Paper presented at the Annual Meeting of the Linguistic Society of America, Portland, OR, January 2012.

Tchekhoff, Claude (1979). 'From Ergative to Accusative in Tongan: An Example of Synchronic Dynamics', in F. Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*. London: Academic Press, 407–418.

te Velde, John R. (2005). *Deriving Coordinate Symmetries: A Phase-based Approach Integrating Select, Merge, Copy and Match*. Amsterdam: John Benjamins.

Terrace, Herbert (1979). *Nim*. New York: Knopf.

Terrill, Angela (2003). *A Grammar of Lavukaleve*. Berlin: Mouton de Gruyter.

Tettamanti, Marco, Alkadhi, Hatem, Moro, Andrea, Perani, Daniela, Kollias, Spyros, and Weniger, Dorothea (2002). 'Neural Correlates for the Acquisition of Natural Language Syntax', *Neuroimage* 17: 700–709.

Tettamanti, Marco, Rotondi, Irene, Perani, Daniela, Scotti, Giuseppe, Fazio, Ferruccio, Cappa, Stefano, and Moro, Andrea (2008). 'Syntax without Language: Neurobiological Evidence for Cross-Domain Syntactic Computations', *Cortex* 45: 825–38.

Thothathiri, Malathi, Kim, Albert, Trueswell, John C., and Thompson-Schill, Sharon L. (2102). 'Parametric Effects of Syntactic–Semantic Conflict in Broca's Area During Sentence Processing', *Brain and Language* 120: 259–264.

Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomioka, Satoshi (2003). 'The Semantics of Japanese Null Pronouns and its Cross-Linguistic Implications', in K. Schwabe and S. Winkler (eds.), *The Interfaces: Deriving and Interpreting Omitted Structures*. Amsterdam: John Benjamins, 321–340.

Tonhauser, Judith (2011). 'Temporal Reference in Paraguayan Guaraní, a Tenseless Language', *Linguistics and Philosophy* 34: 257–303.

Tonhauser, Judith, Beaver, David, Roberts, Craige, and Simons, Mandy (2013). 'Towards a Taxonomy of Projective Content', *Language* 89: 66–109.

Townsend, David, and Bever, Thomas (2001). *Sentence Comprehension*. Cambridge, MA: MIT Press.

Trager, George L. (1955). 'Language', *Encyclopaedia Britannica* 13: 695–702.

Traugott, Elizabeth Closs (2008). 'Grammaticalization, Constructions and the Incremental Development of Language: Suggestions from the Development of Degree Modifiers in English', in Regine Eckardt, Gerhard Jäger, and Tonjes Veenstra, (eds.), *Variation, Selection, Development—Probing the Evolutionary Model of Language Change.* Berlin: Mouton de Gruyter, 219–250.

Travis, Lisa (1989). 'Parameters of Phrase Structure', in Mark R. Baltin and Anthony S. Kroch (eds.), *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press, 263–279.

Traxler, Matthew J., Morris, Robin K. and Seely, Rachel E. (2002). 'Processing Subject and Object Relative Clauses: Evidence from Eye Movements', *Journal of Memory and Language* 47: 69–90.

Trotzke, Andreas (2010). *Rethinking Syntactocentrism: Lessons from Recent Generative Approaches to Pragmatic Properties of Left-periphery-movement in German*. Ph.D. dissertation, University of Freiburg.

Trotzke, Andreas (2012). 'Review of Lieven Danckaert, *Latin Embedded Clauses: The Left Periphery*', *The LINGUIST List* 23. 4367. http://linguistlist.org/issues/23/23-4367.html (Accessed 5th March 2014).

Trotzke, Andreas, and Lahne, Antje (2011). 'Recursion as Derivational Layering: An Amendment to Zwart', *Biolinguistics* 5: 335–346.

Trotzke, Andreas, and Bader, Markus (2013). 'Against Usage-based Approaches to Recursion: The Grammar-Performance Distinction in a Biolinguistic Perspective'. Paper presented at GLOW 36, Lund University, April 2013.

Trotzke, Andreas, Bader, Markus, and Frazier, Lyn (2013). 'Third Factors and the Performance Interface in Language Design', *Biolinguistics* 7: 1–34.

Trubetzkoy, Nikolai (1939). 'Grundzüge der Phonologie', *Travaux du cercle linguistique de Prague* 7.

Trudgill, Peter (1983). *On Dialect: Social and Geographical Perspectives*. Oxford: Blackwell.

Trudgill, Peter (2001). 'Contact and Simplification: Historical Baggage and Directionality in Linguistic Change', *Linguistic Typology* 5: 371–374.

Trudgill, Peter (2004). 'Linguistic and Social Typology: The Austronesian Migrations and Phoneme Inventories', *Linguistic Typology* 8: 305–320.

Trudgill, Peter (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Trueswell, John C., and Kim, Albert E. (1998). 'How to Prune a Garden Path by Nipping It in the Bud: Fast Priming of Verb Argument Structure', *Journal of Memory and Language* 39: 102–123.

Trueswell, John C., Sekerina, Irina, Hill, Nicole M., and Logrip, Marian L. (1999). 'The Kindergarten-Path Effect: Studying On-line Sentence Processing in Young Children', *Cognition* 73: 89–134.

Valenzuela, Pilar M. (1997). *Basic Verb Types and Argument Structures in Shipibo-Konibo*. MA thesis, University of Oregon.

Valenzuela, Pilar M. (2003). *Transitivity in Shipibo-Konibo Grammar*. Ph.D. dissertation, University of Oregon.

Valkama, Kari (2000). *Grammatical Relations in Cebuano*, Helsinki: University of Helsinki.

van Eijk, Jan (1997). *The Lillooet Language: Phonology, Morphology, Syntax.* Vancouver: UBC Press.

van Eijk, Jan, and Williams, Lorna (1981). *Cuystwí Malh Ucwalmícwts: Lillooet Legends and Stories.* Mount Currie: Ts'zil Publishing House.

van Enk, Gerrit J., and de Vries, Lourens (1997). *The Korowai of Irian Jaya: Their Language in its Cultural Context*. Oxford: Oxford University Press.

van Riemsdijk, Henk (1978). A *Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*, Studies in Generative Grammar, 4. Dordrecht: Foris.

van Rossem, Cefas, and van der Voort, Hein (eds.) (1996). *Die Creol Taal, 250 Years of Negerhollands Texts*. Amsterdam: Amsterdam University Press.

Van Valin, Robert, Jr. (1977). *Aspects of Lakhota Syntax*. Ph.D. dissertation, University of California, Berkeley.

Vogel, Ralf (2003). 'Dialectal Variation in German 3-Verb Clusters: Looking for the Best Analysis', *Head Movement and Syntactic Theory: UCLA/Potsdam Working Papers in Linguistics* 10: 199–235.

Vogel, Ralf (2009). 'Skandal im Verbkomplex', *Zeitschrift für Sprachwissenschaft*, 28: 307–346.

von Fintel, Kai (2004). 'Would You Believe It? The King of France is Back! Presuppositions and Truth Value Intuitions', in Marga Reimer and Anne Bezuidenhout (eds.), *Descriptions and Beyond*. Oxford: Oxford University Press, 315–341.

von Fintel, Kai, and Gillies, Anthony (2010). *Must*... Stay... Strong!, *Natural Language Semantics* 18: 351–383.

von Fintel, Kai, and Matthewson, Lisa (2008). 'Universals in Semantics', *The Linguistic Review* 25: 49–111.

Vygotsky, Lev S. (1979/1960). 'The Genesis of Higher Mental Functions', in J. V. Wertsch (ed.), *The Concept of Activity In Soviet Psychology*. New York, M.E. Sharpe, 144–188.

Waddy, Virginia (1889). *Elements of Composition and Rhetoric with Copious Exercises in Both Criticism and Construction.* Cincinnati and New York: Eclectic Press.

Wager, Tor D., and Smith, Edward E. (2003). 'Neuroimaging Studies of Working Memory: A Meta-Analysis', *Cognitive, Affective, and Behavioral Neuroscience* 3: 255–274.

Waldie, Ryan James (2001). *Nuu-chah-nulth Denominal Verbs*. MA thesis, University of Victoria.

Wanner, Eric, and Maratsos, Michael (1978). 'An ATN Approach to Comprehension', in M. Halle, J. Bresnan, and G. A. Miller, (eds.), *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press, 119–161.

Wardlow, Liane, and Ferreira, Victor S. (2008). 'Speaker-External Versus Speaker-Internal Forces on Utterance Form: Do Cognitive Demands Override Threats to Referential Success?', *Journal of Experimental Psychology: Language, Memory, and Cognition* 34: 1466–1481.

Warren, Richard M. (1970). 'Perceptual Restoration of Missing Speech Sounds', *Science* 167: 392–395.

Warren, Tessa, and Gibson, Edward (2002). 'The Influence of Referential Processing on Sentence Complexity', *Cognition* 85: 79–112.

Warren, Tessa, and Gibson, Edward (2005). 'Effects of NP Type in Reading Cleft Sentences in English', *Language and Cognitive Processes* 20: 751–767.

Wasow, Thomas (1997). 'End-Weight from the Speaker's Perspective', *Journal of Psycholinguistic Research* 26: 347–361.

Wasow, Thomas (2002). *Postverbal Behavior*. Stanford: CSLI.

Wasow, Thomas (2009). 'Remarks on Grammatical Weight', *Language Variation and Change* 9: 81–105.

Wasow, Tom, Jaeger, T. Florian, and Orr, David M. (2011). 'Lexical Variation in Relativizer Frequency', in H. Simon and H. Wiese (eds.), *Expecting The Unexpected: Exceptions in Grammar*. Berlin: De Gruyter, 175–195.

Weir, David J. (1988). *Characterizing Mildly Context-Sensitive Grammar Formalisms*. Ph.D. dissertation, University of Pennsylvania.

Wells, Rulon S. (1954). 'Archiving and Language Typology', *International Journal of American Linguistics* 20: 101–107.

Welmers, William Everett (1973). *African Language Structures*, Berkeley: University of California Press.

Westlake, John S. (1908). 'From Alfred to the Conquest', in A.W. Ward and A.R. Waller (eds.), *The Cambridge History of English Literature*. Cambridge: Cambridge University Press. http://www.bartleby.com/211/0703.html (Accessed 5th March 2014).

Whitehead, John (1899). *Grammar and Dictionary of the Babangi Language*, reprinted by the Gregg Press, Ridgewood NJ, 1964.

Whitney, William D. (1897 [1875]). *The Life and Growth of Language: An Outline of Linguistic Science*. New York: D. Appleton.

Wichmann, Søren, Müller, A., Velupillai, V., Wett, A., Brown, C. H., Molochieva, Z., Bishoffberger, J., Holman, E. W., Sauppe, S., Brown, P., Bakker, D., List, J.-M., Egorov, D., Belyaev, O., Urban, M., Hammarström, H., Carrizo, A., Mailhammer, R., Geyer, H., Beck, D., Korovina, E., Epps, P., Valenzuela, P., and Grant, A. (2012). *The ASJP Database (version 15)*. http://email.eva.mpg.de/~wichmann/listss15.zip.

Wickham, Hadley (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer.

Wiechmann, Daniel, and Lohmann, Arne (2013). 'Domain Minimization and Beyond: Modeling PP Ordering', *Language Variation and Change* 25: 1.

Wiklund, Anna-Lena (2007). *The Syntax of Tenselessness*. Berlin: Mouton De Gruyter.

Wilder, Chris, and Gärtner, Hans-Martin (1997). 'Introduction', in C. Wilder, H.-M. Gärtner, and M. Bierwisch (eds.), *The Role of Economy Principles in Linguistic Theory*. Berlin: Akademie Verlag, 1–35.

Wilkins, Matthew R., Seddon, Nathalie, and Safran, Rebecca J. (2012). 'Evolutionary Divergence in Acoustic Signals: Causes and Consequences', *Trends in Ecology and Evolution* 28, 156–166.

Winkler, Susanne (2005). *Ellipsis and Focus in Generative Grammar*. Berlin and New York: Mouton de Gruyter.

Wojdak, Rachel (2005). *The Linearization of Affixes: Evidence from Nuu-chah-nulth*. Ph.D. dissertation, University of British Columbia.

Woll, Bencie, Sutton-Spence, Rachel and Elton, Frances (2001). 'Multilingualism: The Global Approach to Sign Languages', in C. Lucas (ed.), *The Sociolinguistics of Sign Language*. Cambridge: Cambridge University Press, 8–32.

Wurmbrand, Susi (2004). 'West Germanic Verb Clusters: The Empirical Domain', in Katalin É.Kiss and Henk van Riemsdijk (eds.), *Verb Clusters: A Study of Hungarian, German, and Dutch*. Amsterdam: John Benjamins, 43–85.

Wurmbrand, Susi (2005). 'Verb Clusters, Verb Raising, and Restructuring', in Martin Everaert and Henk van Riemsdijk (eds.), *The Syntax Companion*. Oxford: Blackwell, 227–341.

Yanovich, Igor (2012). 'Expressive Complexity in Language: A Mathematical Perspective'. Paper presented at the Conference on Formal Linguistics and the Measurement of Grammatical Complexity, University of Washington, Seattle.

Yngve, Victor H. (1960). 'A Model and an Hypothesis for Language Structure', *Proceedings of the American Philosophical Society* 104: 444–466.

Zeshan, Ulrike (2004). 'Interrogative Constructions in Sign Languages: Cross-Linguistic Perspectives', *Language* 80: 7–39.

Zeshan, Ulrike, Escobedo Delgado, Cesar Ernesto, Dikyuva, Hasan, Panda, Sibaji, and de Vos, Connie (2012). 'Cardinal Numerals in Village Sign Languages—Approaching Cross-Modal Typology', to appear in *Linguistic Typology*.

Zipf, George Kingsley (1935). *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Boston: Houghton Mifflin Company.

Zucchi, Sandro, Neidle, Carol, Geraci, Carlo, Duffy, Quinn and Cecchetto, Carlo (2010). 'Functional Markers in Sign Languages', in D. Brentari (ed.) *Sign Languages*. Cambridge: Cambridge University Press, 197–224.

Zurif, Edgar B. (1995). 'Brain Regions of Relevance to Syntactic Processing', in L. Gleitman and M. Liberman (eds.), *An Invitation to Cognitive Science*. Cambridge, MA: MIT Press, 381–398.

Zwart, C. Jan-Wouter (1995). 'A Note on Verb Clusters in the Stellingwerf Dialect', in Marcel den Dikken and Kees Hengeveld (eds.), *Linguistics in the Netherlands 1995*. Amsterdam: John Benjamins, 215–226.

Zwart, C. Jan-Wouter (1996). 'Verb Clusters in Continental West Germanic Dialects', in James R. Black and Virginia Motapanyane (eds.), *Microparametric Syntax and Dialect Variation*. Amsterdam: John Benjamins, 229–258.

Zwart, Jan-Wouter (2004). 'Unary Merge'. Paper presented at the University of Tilburg, October 28. http://www.let.rug.nl/zwart/docs/h004tilb.pdf (Accessed 5th March 2014).

Zwart, Jan-Wouter (2009a). 'Prospects for Top-down Derivation', *Catalan Journal of Linguistics* 8: 161–187.

Zwart, Jan-Wouter (2009b). 'Uncharted Territory? Towards a Non-cartographic Account of Germanic Syntax', in A. Alexiadou, J. Hankamer, T. McFadden, J. Nuger, and F. Schäfer (eds.), *Advances in Comparative Germanic Syntax*. Amsterdam and Philadelphia: Benjamins, 59–84.

Zwart, Jan-Wouter (2009c). 'Minimalist Construction Grammar'. Paper presented at the TIN-dag, Utrecht, February 7. http://www.let.rug.nl/zwart/docs/h009tin.pdf (Accessed 5th March 2014).

Zwart, Jan-Wouter (2011a). 'Structure and Order: Asymmetric Merge', in C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford: Oxford University Press, 96–118.

Zwart, Jan-Wouter (2011b). 'Recursion in Language: A Layered-derivation Approach', *Biolinguistics* 5: 43–56.

Zwart, Jan-Wouter (2014). 'Top-down Derivation, Recursion, and the Model of Grammar', in A. Trotzke and J. Bayer (eds.), *Syntactic Complexity across Interfaces*. Berlin and New York: Mouton de Gruyter.

Index of languages

Index of names

Index of subjects