

A STOCHASTIC OPTIMALITY THEORY APPROACH TO
SYNTACTIC CHANGE

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Brady Z. Clark
July 2004

© Copyright by Brady Z. Clark 2004
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Paul Kiparsky
(Principal Co-Advisor)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Elizabeth Traugott
(Principal Co-Advisor)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Joan Bresnan

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Peter Sells

Approved for the University Committee on Graduate Studies:

Abstract

This thesis challenges the claim that syntactic change should be modeled in terms of adjustments in the probability distribution of competing grammars (Kroch 1989b; Yang 2000). Building on work of Boersma (1998), Boersma and Hayes (2001), and Bresnan et al. (2001), I propose a novel formal framework for modeling syntactic change that is firmly embedded in the inherent variability tradition (Weinreich et al. 1968). In this approach, grammars are conceived of as constraint rankings on a continuous scale of real numbers. Further, in the process of speaking or hearing (i.e., at any evaluation point) the rank of each constraint is slightly perturbed by adding a random value drawn from a normal distribution. This perturbation of constraint ranking leads to intraspeaker variation and, consequently, defines a probability distribution over outputs from the grammar for a given input, where, roughly speaking, the input is the meaning of an utterance or the speech stream. Given this picture of linguistic competence, syntactic change is reflected by adjustments in constraint strength along a continuous scale of real numbers.

Two case studies on syntactic change in early English form the empirical core of this thesis: the gradual decline of right-headed structures (e.g., OV), and change and variation in the syntax of subjects. These case studies draw on data from secondary sources, as well as novel evidence from the York-Toronto-Helsinki Parsed Corpus of Old English Prose and the Penn-Helsinki Parsed Corpus of Middle English, Second Edition. The analyses proposed capture generalizations about English clause structure change discussed in earlier work, with additional empirical benefits. Certain logically possible but unattested word orders in early English are predicted to be absent. Further, a theory embedded in this framework allows one to give a unified analysis of change and variation in the syntax of different types of subjects in early English, accounting for why subject pronouns gradually adopt the syntax of full noun phrase subjects in Middle English.

Apart from these empirical gains, the case studies highlight several important properties of the grammatical model advocated for in this thesis. First, the grammatical model contains a nondeterministic, noncategorical, and quantitative component that captures intratextual variability displayed by both of the phenomena explored in the case studies. Second, the steady quantitative rise in the rate of use of an innovation is reflected by gradual adjustments of constraint strength. Third, the model accounts directly for the observation that categorical phenomena at one stage of a language show up as statistical preferences at another.

Acknowledgments

The seed of inspiration for this work came from a series of papers by Joan Bresnan and collaborators applying the framework of Stochastic Optimality Theory to problems in syntactic variation. I saw in that work a new grammatical model for syntactic change that overcame some shortcomings of previous approaches, as well as an opportunity to investigate in one work several interests I had been cultivating for years; e.g., the formal foundations of syntactic theory and word order variation in early English. This project began in 2001, progressed in fits and starts, and then gelled into this document in early 2004.

I would never have completed this thesis without the aid, financial and otherwise, of many Stanford faculty and staff, both at the Department of Linguistics and at CSLI. I owe an enormous debt of gratitude to David Beaver for many reasons, both academic and nonacademic, but most especially for hiring me back in 1999 to work on a project that has never ceased to yield discoveries. Our work on focus and focus sensitivity has been the most productive of my graduate school career. Likewise, many thanks to Stanley Peters for presenting me with an employment opportunity in 2000 that has subsequently taken me from Verona to the interior of a Naval ship, as well as taught me many skills that I would not have learned otherwise. Finally, it would have been impossible to balance work, family, and bureaucracy without all of the staff members that have made my life easier over the last seven years. I would like to single out Anita Black, Laura Burns-Wood, Christina Doering, Dorothy Hodges, Melanie Levin, Natalie Mendoza, Natasha Newson, Cristy Osborne, Emma Pease, Coco Relova, Mary Reyes (R.I.P.), Najwa Salame, Allen Sciutto, Julie Shih, Trudy Vizmanos, and Gina Wein. Thank you!

My parents saved us countless times from financial ruin and despair. I am forever grateful for their unwavering love and support, even when my recklessness was to blame for our occasional descents into bankruptcy. Thank you!

I owe an immeasurable debt to my thesis committee for their unstinting commitment to this project. Paul Kiparsky is a constant source of inspiration. His insight on all aspects of historical linguistics and his attention to detail challenged me to rethink and expand the boundaries of this work up until the end. Elizabeth Traugott has been, from the very beginning of my time at Stanford, the perfect mentor. Elizabeth provided heaps of inspiration and invaluable guidance as I explored early English syntax. Both Paul and Elizabeth taught me to how to think about historical syntax. I will miss our weekly conversations

greatly. Joan Bresnan has been constantly supportive of this work, even during long periods of drought. Many of the really good ideas in this thesis have their roots in meetings with Joan. Peter Sells started me thinking about the foundations of Optimality-Theoretic syntax in a directed reading way back in the summer of 1999. In many ways, this thesis is my attempt at answering some of the questions we explored half a decade ago. Further, this thesis would not have been possible without Peter's groundbreaking work on Swedish clause structure. That said, the significance of what all four of my committee members have contributed to my development as a linguist far exceeds the content of this thesis. Thank you!

This thesis has benefited from input from a wide community of scholars. Various parts have been presented at the Spring 2001 Berkeley-Stanford Workshop on Historical Morphosyntax at UC Berkeley, at the January 2002 Annual Meeting of the Linguistic Society of America in San Francisco, at Studies in the History of the English Language (SHEL) II in March 2002 at the University of Washington, at New Ways of Analyzing Variation (NWAV) 31 at Stanford University in October 2002, and at Diachronic Generative Syntax (DIGS) VIII in June 2004 at Yale University. I have also presented aspects of this thesis in courses and workshops at Stanford University. Gerhard Jäger and Christopher Potts sent me detailed comments that greatly improved the content of Chapter 2. Thank you!

This thesis is dedicated to Gina, Malachi, and Riley. You are my heart. We don't have much time — let's turn the world into a holiday.

Contents

1	Introduction	1
1.1	Change and variation	2
1.2	The grammatical model	6
1.3	Case studies	8
1.4	Methodological issues	10
1.4.1	Dialect and contact issues	10
1.4.2	Genre issues	11
1.4.3	Corpus issues	11
1.5	Overview of the thesis	12
2	A model-theoretic characterization of LFG	13
2.1	Introduction	14
2.2	Model-theoretic syntax	15
2.3	Lexical-Functional Grammar	16
2.3.1	C-structure	17
2.3.2	Constraints on c-structure	26
2.3.3	Well-formed formulae	27
2.3.4	Satisfaction definition	30
2.3.5	F-structure	32
2.3.6	The correspondence mapping	35
2.3.7	Advantages of the model-theoretic approach to syntax	38
2.3.8	Constraints on LFG models: G_{inviol}	41
2.3.9	In sum	52
2.4	The Architecture of OT-LFG	52

2.4.1	Overview of architecture and main hypotheses	52
2.4.2	An example: Verb positioning in English	55
2.4.3	Components of OT-LFG grammars	62
2.4.4	In sum	76
2.5	Stochastic Optimality Theory	76
2.5.1	Stochastic OT-LFG	77
2.5.2	Example: StOT-LFG approach to verb positioning	82
2.6	Closing remarks	88
3	The decline of right-headed clauses in early English	91
3.1	Clause structure variation in early English	92
3.1.1	Evidence for clause structure variation in Old English	96
3.1.2	Evidence for clause structure variation in early Middle English	98
3.1.3	Exocentric S in early English	100
3.2	A Stochastic OT-LFG analysis	105
3.2.1	A theory of clause structure variation	105
3.2.2	Modeling change in clausal headedness	126
3.3	Closing remarks	133
3.3.1	Directionality of change	134
3.3.2	The Constant Rate Effect	150
4	The syntax of subjects in early English	161
4.1	Introduction	162
4.2	The formal expression of pronominal content	164
4.3	Change and variation in the syntax of subjects in early English	167
4.3.1	Subjects and topicalization	170
4.3.2	Subjects and negative adverbs	184
4.3.3	Evidence that subject pronouns did not project	195
4.3.4	Evidence that subject pronouns were phonologically dependent	198
4.3.5	In sum	199
4.4	A Stochastic OT-LFG analysis	202
4.4.1	A theory of subjects	202
4.4.2	Diachrony in the syntax of early English subjects	231
4.4.3	Further implications: Null subjects in early English	245

4.5	Conclusion	249
4.5.1	Positional discrepancies in related languages	250
4.5.2	Cause(s) of change in the syntax of subjects in Middle English	251
5	Conclusion	253
5.1	Introduction	253
5.2	Summing up	254
5.3	On stochastic grammar	258
5.3.1	Methodological issues	258
5.3.2	Extragrammatical factors determine statistical preferences	263
5.3.3	The multiplicity of genres	269
5.3.4	The reification of style	271
5.3.5	In sum	274
5.4	Future prospects	274
5.4.1	Extending the formal specification of StOT-LFG	274
5.4.2	Extending the case studies	276
5.4.3	Open questions about change	278
	Bibliography	282
	Author Index	314

List of Tables

3.1	Violations incurred by main and subordinate clauses	125
3.2	Stages of learning <i>Ælfric's Lives of Saints</i>	156

List of Figures

1.1	The idealized timecourse of linguistic change	4
2.1	Categorical constraint ranking with ranges of variation:	78
2.2	Free constraint ranking with ranges of variation:	78
2.3	Reversal of constraint dominance:	79
2.4	The position of <i>never</i> with respect to the main verb (Ellegård 1953, 184) . . .	84
2.5	Partial stochastic grammar (1425-1475) — verb positioning	85
2.6	GLA results for 1425-1475 input	85
2.7	Partial stochastic grammar (1500-1525) — verb positioning	86
2.8	GLA results for 1500-1525 input	86
2.9	Partial stochastic grammar (1575-1600) — verb positioning	86
2.10	GLA results for 1575-1600 input	87
2.11	Partial stochastic grammar (<i>faux</i>) — verb positioning	87
2.12	GLA results for <i>faux</i> input	88
2.13	A mismatch between the learner’s form and the adult form	89
2.14	The learning step: adjusting the ranking values	89
2.15	Training schedule	90
3.1	Estimated frequencies of structures (3.1a-c) in Chron A, Scribe 1 and <i>Festis Marie</i>	127
3.2	Ranking values for Chron A, Scribe 1	128
3.3	Partial stochastic grammar for late Old English (Chron A, Scribe 1)	129
3.4	GLA results for OE input	130
3.5	Ranking values for <i>Festis Marie</i>	131
3.6	Partial stochastic grammar for early Middle English (<i>Festis Marie</i>)	132
3.7	GLA results for EME input	133

3.8	Frequency of INFL-medial vs. INFL-final in Old English main and subordinate clauses (Pintzuk 1999: 208); N = total number of clauses analyzed . . .	152
3.9	Slope and input probability parameters of the frequency of INFL-medial structure in Old English main and subordinate clauses, separate VARBRUL runs (Pintzuk 1999:222)	153
3.10	Violations incurred by main and subordinate clauses	154
3.11	Results of learning Old English frequency data	158
3.12	<i>Anti-Ælfric's Lives of Saints</i>	159
4.1	The projection/dependence matrix	165
4.2	Rate of inversion of full noun phrase subjects after objects (Koopman 1998, 136)	175
4.3	Postverbal subjects and negative elements in OE (Helsinki Corpus, data from Rissanen 1999, 191)	187
4.4	Postverbal subjects and negative elements in EME (Helsinki Corpus, data from Rissanen 1999, 192)	187
4.5	Postverbal subjects and negative elements in LME (Helsinki Corpus, data from Rissanen 1999, 193)	188
4.6	Postverbal subjects and <i>not</i> in early Modern English (Helsinki Corpus, data from Rissanen 1999, 196)	190
4.7	Post-aux subjects and <i>not</i> in early Modern English (Helsinki Corpus, data from Rissanen 1999, 196)	190
4.8	Estimated frequencies of structures (4.100)-(4.105) in early English texts . . .	235
4.9	Partial stochastic grammar of <i>Ælfric's Lives of Saints</i>	237
4.10	Partial stochastic grammar for OE (<i>Ælfric's Lives of Saints</i>)	238
4.11	GLA results for <i>Ælfric's Lives of Saints</i> input	239
4.12	Partial stochastic grammar of <i>Vices and Virtues</i>	240
4.13	Partial stochastic grammar for EME (<i>Vices and Virtues</i>)	241
4.14	GLA results for <i>Vices and Virtues</i> input	242
4.15	Partial stochastic grammar of <i>Wycliffite Sermons</i>	243
4.16	Partial stochastic grammar for LME (<i>Wycliffite Sermons</i>)	243
4.17	GLA results for <i>Wycliffite Sermons</i> input	244