

Book proposal for:
Trends in Linguistics. Studies and Monographs (TiLSM)

Author: Daniel Wiechmann

1 Title

A corpus-based view on the processing of English relative clauses (working title)

2 Description of the topic

Recent years have witnessed a resurfacing of interest in the relationship between linguistic knowledge and language use. As a matter of course, the range of views encompasses virtually all conceivable positions, but among those linguists that address theoretical issues through the analysis of naturalistic language data, a *performance-* or *usage-based* (UB) view is becoming ever more popular. On this a view, grammar is the cognitive organization of one's experience with language and as such heavily affected by the frequency with which linguistic expressions have been perceived. To this day UB-theories have documented frequency effects at various areas of linguistic interest, from language acquisition to diachronic change. One particularly interesting area to investigate the relationship between grammar and usage is adult language comprehension, where the amount of available data (from both psycholinguistic experimentation and linguistic corpora) is by far most extensive. In the domain of lexical processing, the UB-perspective on the role of frequency information is corroborated by an abundant amount of psycholinguistic research demonstrating a strong negative correlative relationship between the processing demand of a lexical expression and its frequency in the ambient language. Frequent words are accessed faster than rare words as evidenced by faster lexical decision times in comprehension and by faster naming times in production. However, syntactic processing is often conceived of as being categorically different from lexical processing and the role of usage frequency information is still unclear for this domain. Many psycholinguistic theories follow modern generative theories of language and assume a bipartite architecture of grammar consisting (roughly) of words and rules. In line with this ontological distinction, these accounts premise that lexical items are stored in and accessed from a mental lexicon, whereas syntactic structures are the product of some kind of rule application. In consequence, lexical and syntactic processing are presumed to be governed by different sets of cognitive mechanisms. UB-accounts, on the other hand, typically gravitate towards a constructionist architecture, which discards a principled distinction between lexicon and grammar and instead characterizes linguistic knowledge as a large assembly of symbolic structures, so called constructions. The present study offers a usage-based account of how humans comprehend complex linguistic structures. It proposes a theory of constructional access, which treats syntactic patterns as complex and abstract signs. In this view, syntactic processing is subject to the very same dynamics as lexical processing and should yield the same type of frequency effects. On the basis of a comprehensive corpus-based analysis of English relative clause constructions, the study then pursues the hypothesis that the processing demand of an arbitrary construction is a function of the degree of cognitive entrenchment of that construction in a language user's mental grammar.

3 Contribution made to current debate

The analysis departs from more traditional approaches to language processing, in so far as the latter typically assume that it is a set of intrinsic properties of a construction type (e.g. the complexity of that construction) that governs its processing difficulty. While architectural constraints from language production may govern certain linguistic choices under real-time pressure to some extent, it is argued here that it is more felicitous to situate the effects of complexity at the social level, which embodies processes of conventionalization. The view presented here thus acknowledges an important connection between grammar and usage but suggests that the causal powers associated with the intrinsic properties of linguistic patterns first and foremost figure in the shaping of grammars over historical time, rather than in on-line processing.

4 Proposed structure of the book

The book is divided into six parts. Parts 1 and 2 are dedicated to setting the stage and explicating all relevant theoretical constructs. These include conceptions from recent linguistic and psycholinguistic theorizing, most notably conceptions from construction grammar and exemplar theory. Chapter 3 presents the corpus data used in the study and investigates the data in regard to a number of variables that have figured in previous explanatory approaches to processing difficulty. Chapter 4 moves from bivariate to multifactorial statistical models and discusses why the employed analytical tools have been used and how. Chapter 5 discusses the results of the study and Chapter 6 concludes it. A more detailed sketch of the proposed structure is given below:

1. Introduction
 - 1.1. Overture: Some precursors and some prerequisites
 - 1.1.1. Symbolization and mental states
 - 1.1.2. Signaling processing instructions I: form to meaning
 - 1.1.3. Signaling processing instructions II: form to form
 - 1.1.4. Conventional patterns as routinized instructions
 - 1.2. Summary
2. Towards a theoretical framework of the right kind
 - 2.1. The merits of being sign-based
 - 2.1.1. Regularity in language: rules and schemas
 - 2.1.2. Uniform representation of linguistic knowledge
 - 2.2. The merits of being usage-based
 - 2.2.1. Effects of frequency
 - 2.3. Language processing as Constructional Access
 - 2.3.1. Memory-based language processing
 - 2.3.2. Categorizing complex constructions
 - 2.4. Summary

3. Methods, data, and beyond
 - 3.1. Corpus data used in the analysis
 - 3.2. A roadmap for the analysis
 - 3.3. Head features
 - 3.3.1. Morphosyntactic realization of the head
 - 3.3.2. Definiteness of the head
 - 3.3.3. Contentfulness of the head
 - 3.3.4. Animacy of the head
 - 3.3.5. Concreteness of the head
 - 3.4. Features of the relative clause
 - 3.4.1. Finiteness
 - 3.4.2. Transitivity
 - 3.4.3. Relativized role
 - 3.5. Features of the main clause
 - 3.5.1. Transitivity
 - 3.5.2. External role and type of embedding
 - 3.6. Cross-clausal features
 - 3.6.1. Transitivity configurations
 - 3.6.2. Syntactic parallelism
 - 3.6.3. Thematic parallelism
 - 3.6.4. Interference and discourse-function
 - 3.7. Comparison with findings from different corpora
 - 3.8. Summary
4. Expanding horizons: RCC in ambient configurational space
 - 4.1. Non-finite RCCs - bivariate prelude
 - 4.2. A configural perspective on non-_nite RCCs
 - 4.2.1. Mining association rules via k-optimal patterns analysis
 - 4.2.2. Detecting patterns via Configural frequency analysis
 - 4.2.3. Identifying exemplar clusters
 - 4.3. Finite RCC
 - 4.3.1. Finite subject RCCs
 - 4.3.2. Finite non-subject RCCs
 - 4.4. Constructional schemas and relativizer omission
 - 4.5. Summary
5. General discussion
6. Conclusion and outlook

5 Brief narrative CV of the author

Daniel Wiechmann studied English linguistics, cognitive psychology and theoretical philosophy at the Universities of Marburg, Stellenbosch (South Africa) and Hamburg. In 2005, he joined the English department at the University of Jena, where he received his PhD (Dr. phil) in 2010. He is currently employed as a research fellow in the DFG-funded research project "*A comprehensive survey of major contrasts between English and German*".

6 Description of target audience

Cognitive linguists, corpus linguists, psycholinguists.

6 List of competitor publications

Even though relative clauses play a pivotal role in research into syntactic complexity and sentence processing, there is no comparable book-length corpus-study on the market to this day. However, there are a number of studies that may help classify the present proposal as they share crucial theoretical and/or methodological commitments. These include (but are not limited to):

Diessel, H. 2004. *The Acquisition of Complex Sentences*. [Cambridge Studies in Linguistics 105]. Cambridge: Cambridge University Press.

Gries, S. Th. 2003. *Multifactorial analysis in corpus linguistics: a study of Particle Placement*. London & New York: Continuum Press.

Wulff, S. 2009. *Rethinking Idiomaticity: A Usage-based Approach*. London/New York: Continuum Press.

8 Approximate length

300 pages