

The Syntactic Process

Mark Steedman

(University of Edinburgh)

Cambridge, MA: The MIT Press
(Language, speech, and communication series), 2000, xiv+330 pp; hardbound, ISBN 0-262-19420-1, \$35.00

Reviewed by
Joakim Nivre
Växjö University

The main asset of categorial grammar has traditionally been the close association between syntactic description on the one hand and a transparent and well-founded semantics on the other. Consequently, categorial grammar has primarily been favored by semantically oriented theorists and has arguably played an important role in the development of formal semantics for natural languages (see, for example, Lewis [1970] and Montague [1973]). From the point of view of grammatical theory, however, frameworks based on categorial grammar have usually been problematic either by being too restricted in their expressive power or by not being restricted enough. Thus, classical categorial grammar, where the only permissible operation is functional application, is equivalent in expressive power to context-free phrase structure grammar and is therefore considered by most researchers as too restrictive. By contrast, the various extensions sometimes grouped under the heading of “flexible categorial grammar,” which allow operations such as functional composition and type raising, have been criticized for being too permissive. In the most extreme case, these extensions may result in permutation-completeness, meaning that any grammar will generate the permutation closure of the context-free language defined by the lexicon (van Benthem 1986; Moortgat 1989). But even if this can be avoided, it is normally the case in these frameworks that one and the same string has multiple derivations resulting in the same semantic representation, a phenomenon which is often referred to as “spurious ambiguity.”

Mark Steedman, in his book *The Syntactic Process*, presents the framework of combinatory categorial grammar (CCG), which attempts to steer a steady course between the Scylla of context-freeness and the Charybdis of permutation-completeness, using a suitably constrained set of operators. Moreover, Steedman argues that the alternative derivations, far from being spurious, are needed to capture aspects of linguistic structure that have traditionally been neglected in grammatical theory, notably coordination, parentheticals and, last but not least, intonation structure. This leads to a new conception of syntactic structure, which is no longer seen as an independent level of representation but rather as a trace of the process by which an interpretation is derived. The claim is that different derivations of the same predicate-argument structure actually correspond to different ways of structuring a content, something that is often reflected in the intonation structure of utterances. Steedman therefore proposes the term **information structure** for this notion of syntactic structure, which is claimed to unify the separate notions of surface structure and intonation structure postulated in other theories of grammar (see, for example, Selkirk [1984]). Another advantage of this unorthodox notion of constituent structure, according to the author, is that it directly supports incremental processing in such a way that every processing step cor-

responds to the construction of a complete constituent with a well-defined semantic interpretation.

The book consists of 10 chapters, with Chapters 2–9 grouped into three parts and Chapter 1 as a general introduction. Part I (Chapters 2–5) is entitled “Grammar and Information Structure” and serves as a general introduction to CCG while at the same time developing the idea of syntactic structure as information structure. Chapter 2 situates the approach within the broader tradition of generative grammar. Chapter 3 introduces the technical machinery of CCG with motivating linguistic examples, and Chapter 4 defines the space of possible CCGs more exactly. These three chapters lead up to Chapter 5, where the argument is made in favor of an unorthodox derivation-based notion of syntactic structure capturing both information structure and intonation structure. Part II, “Coordination and Word Order,” consists of two connected case studies, the first dealing with the notorious cross-serial dependencies in Dutch (Chapter 6), and the second dealing with gapping in English and Dutch (Chapter 7). Part III, “Computation and Performance,” is devoted to the processing of CCG and mainly deals with computational issues, although human linguistic performance is also touched upon. Chapter 8 examines questions of expressive power, while Chapter 9 outlines a specific parsing architecture and makes the argument that CCGs, unlike most other competence grammars, can be used directly in semantically incremental parsing. Chapter 10, finally, summarizes the architecture of the theory as a whole and briefly discusses its role in acquisition and performance.

This book will be of interest to theoretical and computational linguists alike, although its strength lies clearly on the theoretical side. It is true that most of Part III is devoted to the computational processing of CCG, but despite many interesting ideas, this part of the book remains somewhat inconclusive. For example, we are told that CCGs can be parsed in polynomial time with a worst-case complexity of $O(n^6)$, but this depends on results for linear indexed grammars by Vijay-Shanker and Weir (1990, 1994) and an assumed equivalence between CCG and linear indexed grammars, which in turn depends on a certain way of interpreting the rule schemata for type raising in CCG. Similarly, the proposed parsing architecture involving a breadth-first incremental parser combined with contextual disambiguation in order to reduce the amount of nondeterminism is interesting but still rather sketchy. And Part II, as pointed out by the author himself, is a self-contained and purely linguistic monograph, which is probably of less interest to most computationally oriented readers. I nevertheless claim that this is a book that deserves serious attention from both computational and theoretical linguists, mainly on account of the material in Part I. As noted earlier, these four chapters serve a dual purpose. On the one hand, they form the introduction to CCG that, at least according to one of the quotes on the dust jacket, the community has been waiting for. On the other hand, they contain a very interesting theoretical proposal concerning the nature of syntactic structure, a proposal that not only integrates phonology, syntax, semantics, and pragmatics but also takes computational aspects into account. In this way, Steedman’s book demonstrates very clearly that categorial grammar in general and CCG in particular is one of the most promising approaches available, both in the field of grammatical theory and in the domain of grammar-based natural language processing.

References

Lewis, David. 1970. General semantics. *Synthese*, 22: 18–67.
Montague, Richard. 1973. The proper

treatment of quantification in ordinary English. In J. Hintikka, J. M. E. Moravcsik, and P. Suppes, editors, *Approaches to Natural Language: Proceedings*

- of the 1970 Stanford Workshop on Grammar and Semantics*. Reidel, Dordrecht, pages 221–242.
- Moortgat, Michael. 1989. *Categorial Investigations*. Foris, Dordrecht.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax*. The MIT Press, Cambridge, MA.
- van Benthem, Johan. 1986. *Essays in Logical Semantics*. Reidel, Dordrecht.
- Vijay-Shanker, K. and David Weir. 1990. Polynomial time parsing of combinatory categorial grammars. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*. Pittsburgh, PA, pages 1–8.
- Vijay-Shanker, K. and David Weir. 1994. The equivalence of four extensions of context-free grammar. *Mathematical Systems Theory* 27: 511–546.

Joakim Nivre is assistant professor at the School of Mathematics and Systems Engineering, Växjö University. His research interests include formal syntax and semantics, corpus linguistics, and natural language processing. Nivre's address is School of Mathematics and Systems Engineering, Växjö University SE 351 95 Växjö, Sweden; e-mail: Joakim.Nivre@msi.vxu.se; URL: <http://www.masda.vxu.se/~nivre>.