

The last two chapters of the book form Part IV, subtitled "Explorations." The first of these comes as something of an agreeable surprise, and is certainly the best chapter in the book. The authors, Hauenschild and Busemann, investigate the possibility of adapting GPSG to MT. In particular, they address the problem of developing a "constructive" (or constructional) version of the "purely axiomatic" version found in Gazdar et al. (1985). Here at last is something of an answer to the very apposite question posed by Kimmo Kettunen (1986) in this journal.

The second chapter might have been a similar exploration of LFG by Schmidt. Instead, it "tries to give an idea of how to overcome some weaknesses of [EUROTRA's] CAT-formalism . . . by relating it to LFG" (p. 239). The last two sentences of the book seem to confirm the impression that perhaps they might have done better just to start with LFG in the first place:

This chapter was a glimpse at a theory from which a great deal has been imported into CAT, namely LFG. However, some of the virtues of the original have been ignored, above all the most clever LFG mechanism, the functional uncertainty device. (p. 250)

#### REFERENCES

- Bresnan, J. (ed.) 1982 *The Mental Representation of Grammatical Relations*. Cambridge, MIT Press, Cambridge, MA.  
 Gazdar, G.; Klein, E.; Pullum, G. K.; and Sag, I. 1985 *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford.  
 Kettunen, K. 1986 "Is MT Linguistics?" *Computational Linguistics*, 12(1), 37-38.

*Harold Somers*, a lecturer at UMIST, is a member of the machine translation research group there with a special interest in the relationship between syntax and semantics. Although formerly involved in the EUROTRA project, he has not been associated with that project for three years. Somers's address is: Centre for Computational Linguistics, UMIST, P.O. Box 88, Manchester M60 1QD, U.K. E-mail: hls@ccl.umist.ac.uk

#### STUDIES IN COMPUTER-AIDED LEXICOLOGY

**Martin Gellerstam (ed.)**  
 (University of Göteborg)

Stockholm: Almqvist & Wiksell International, 1988,  
 375pp. (Data Linguistica 18)  
 Hardbound, ISSN 0347-948X, ISBN 91-22-01258-3,  
 SEK287.00

*Reviewed by*  
*Martha Evens*  
*Illinois Institute of Technology*

This volume is a collection of papers in honor of the sixtieth birthday of Sture Allén, Chairman of the Department of Computational Linguistics at the University of Gothenburg and Permanent Secretary of the Swedish Academy, written by his students and colleagues. These papers appear

in excellent English but they describe work on Swedish (although many of the authors represented here have also worked on problems in English lexicology). The organization of the book, in alphabetical order by last name of author, has the effect of emphasizing the amazing breadth of Allén's work, as we move from a paper on applications of lexical databases in linguistics to a paper on the creation of such a database, or from a paper on the structure of the lexicon to a paper on how a computer program can access a lexical database.

The problem of creating a lexical database from dictionary data is discussed in three papers. Rudolf Rydstedt gives a preliminary report on the construction of the lexical database for the *Dictionary of the Swedish Academy*, a project headed by Allén himself (comparable to the construction of the new *Oxford English Dictionary* by the University of Waterloo and IBM in cooperation with the Oxford University Press). The paper includes a description of the languages used for defining the data, the problems of structuring fixed- and free-format text, and the question of creating a database that can be distributed in CD-ROM form. Sture Berg and Kaisa Samuelson exhibit the cooperation between linguist and computer scientist that is such a happy feature of this book in a description of the production of a glossary for use in spelling correction; the major problems are the generation of all inflected forms and the analysis of the novel compounds that are a common feature of modern Swedish. Rolf Gavare talks about the problem of sorting in alphabetical order any collection of "words" including uppercase letters, numerals, Greek and other non-Roman characters, and logograms like percent signs and ampersands.

A number of the papers discuss the structure of the lexical database and desirable features of the lexical entry. Staffan Hellberg proposes a novel classification of adjectives. Martin Gellerstam and Maria Toporowska-Gronostaj both discuss methods of representing the way verbs and their arguments combine. Bo Ralph looks at semantic rather than syntactic structures for verbs, concentrating mainly on taxonomy.

Some lexical databases are created primarily for use by other computer software; others are intended to be used by human beings. Anna Sägval Hein describes the lexical database used by her parser for Swedish, a unification-based chart parser designed for such diverse sublanguages as news agency telegrams, medical texts, and dictionary definitions. Christian Sjögreen describes a project, also designed by Allén, to create a commercial dictionary from a lexical database, complete to the writing of the printer's tape. Guðrún Magnúsdóttir discusses problems of accessing the lexicon in the machine translation process. The advantages of a lexical database as a source for linguistic research are illustrated by studies of synaesthesia (Åsa Abelin), semantic change (Birgitta Ernby), and loan words (Kerstin Norén).

Two papers of extraordinary interest span all these categories. Jerker Järborg presents a formal structure for the description of both syntactic and semantic features. Sven-

Göran Malmgren describes regularities in polysemy: types of metaphoric transfer of meaning in nouns, regular extensions in verb meanings, and changes in adjective meanings as the argument changes. While these papers are written about Swedish, they make one itch to try out these ideas on an English corpus.

*Martha Evens* is a professor of Computer Science at Illinois Institute of Technology. She is a co-author of one book on lexical-semantic relations and the editor of another. In 1984 she was president of the Association for Computational Linguistics. Evens's address is: Department of Computer Science, Illinois Institute of Technology, 10 West 31st Street, Chicago, IL 60616. E-mail: csevens@iitvax.bitnet

## COMPUTATIONAL LINGUISTICS: AN INTRODUCTION

**Ralph Grishman**

(New York University)

Cambridge, U.K.: Cambridge University Press, 1987, viii + 193 pp.

(Studies in Natural Language Processing)

Hardbound, ISBN 0-521-32501-1, \$39.50; Softbound, ISBN 0-521-31038-5, \$14.95 (20% discount to ACL members)

*Reviewed by*

*John Roach*

*Virginia Polytechnic Institute and State University*

Grishman's book provides an introductory overview of the computational linguistics field. Instructors looking for a good introductory text to use in a course devoted to computational linguistics should consider this book as a strong candidate if they wish to emphasize a linguistics approach. The book consists of five sections: introduction (5%), syntax analysis (45%), semantic analysis (29%), discourse analysis and information structuring (11%), and language generation (7%), as well as a short section of 34 exercises covering all chapters of the book. The book's emphasis reflects the predominant trends in the field. Grishman avoids extreme positions and gives a remarkably neutral account. By the same token, however, students will hardly know anything of the controversies that have shaken the field.

Many different approaches could be taken when attempting a broad survey of computational linguistics. Artificial intelligence could be emphasized (including accounts of past programs), a cognitive science approach might be attempted, or a linguistic orientation might be the main emphasis. Grishman elects the linguistic approach, although some artificial intelligence programs are discussed. The book *Natural Language Processing* by Harry Tennant (1981), which covers many AI programs, is referenced (together with Winograd's *Language as a Cognitive Process* [1983]) as a supplement in the introductory chapter. Using both of these books together could make for a very good course: Grishman's book provides a strong linguistic orientation, and Tennant's book presents practice and appli-

cations. The main difficulty with using both texts together would be the need to integrate the disparate approaches. By taking a linguistic viewpoint, Grishman's text fails fully to cover many different computational approaches to language. For use in a class, the book must be supplemented.

As an example of topics not (sufficiently) covered, consider case frame analysis. The text barely mentions case frames, and the only analysis of canonical verbs with agents, instruments, affected entity, etc., occurs in a brief discussion of Schank's conceptual dependency theory. Any instructor wishing to cover the large amount of computational linguistics research in this area will have to draw more material from elsewhere.

Grishman does concentrate on the central themes of linguistics: transformational grammar, translating surface forms into semantic meaning structures (first-order logic statements in this case) and, to a lesser extent, discourse analysis.

In the syntax chapter, Grishman first asks whether syntactic analysis is really necessary, since some researchers today question the need. He observes that syntactic analysis is the norm and highly useful. He also briefly addresses the difficulty of parsing non-well-formed inputs. He then presents the Chomsky hierarchy of grammars, regular, context-free, and context-sensitive, followed by a rather nice example of a context-free grammar and a parser for it. Sections on transformational grammar and an associated parser are covered in thorough detail. Augmented phrase-structure grammars (including NYU's Linguistic String Project and ATNs) are also covered. The syntax chapter concludes with brief discussions of some specific problems: ambiguities in attaching prepositional phrases to what they should modify and analyzing coordinate conjunctions. The latest advances in syntactic theory (government-binding, X-bar syntax, etc.) have not been included. Nevertheless, this is a very full chapter that provides quite good coverage of syntactic parsing techniques.

The semantic analysis chapter introduces predicate logic as the formal language into which surface forms can be translated. Grishman also discusses shortcomings of predicate logic as a representation language, after which he mentions Montague semantics but provides no treatment.

Semantic constraints, semantic grammars, and sublanguages are explained briefly but well. Grishman then discusses some classic semantic problems, including anaphora and reference (a very nice treatment), discourse entities, and noun phrases. The chapter concludes with an interesting discussion of uses of logic as a representation language, including deducing consequences from sentences using theorem provers. As with the syntax chapter, the latest hot buttons in semantics are not treated (situation semantics, for example, is not mentioned anywhere) but the coverage of topics presented contains insightful discussions.

Examples are well chosen and informative. Occasionally, the reader might wish for more extensive examples.

The syntax and semantics chapters occupy much of the book. The chapter on discourse and structured representations is therefore rather short. Grishman reviews what we