

that the only objective of the book is description of linguistic phenomena. It is difficult to imagine practical application of the theory in natural-language processing systems. Although the book was not aimed at presenting practical applications, a short chapter on this topic could dispel the doubts of a reader studying this type of problem for the first time. Furthermore, the addition of an index would have facilitated the reader in returning to certain issues or unmemorized definitions.

Still, this book is one of the more interesting recent publications on the application of logic in natural-language description. Reading it will inspire further research on the logical structure of natural language, and is highly recommended.

*Leonard Bolc* is the editor of several collections of papers on various aspects of natural-language systems, including the recent *Natural-language parsing systems* (Springer-Verlag). His address is: Instytut Podstaw Informatyka, Polskiej Akademii Nauk, PKiN, pok. 1050, 00-901 Warszawa, Poland.

---

**MEMORY AND CONTEXT FOR LANGUAGE  
INTERPRETATION  
(STUDIES IN NATURAL-LANGUAGE PROCESSING)**

**Hiyan Alshawi**

(SRI International, Cambridge, England)  
Cambridge, England: Cambridge University Press,  
1987, ix + 188 pp.  
ISBN 0-521-34059-4, \$29.95 (hb) (20% discount to  
ACL members)

*Reviewed by*  
*Jean-Pierre Corriveau*  
*University of Toronto and*  
*Bell-Northern Research*

This book is a reorganization of a 1983 doctoral thesis. It seems little effort has been spent to take into account the impressive amount of research in text comprehension since 1983. Indeed, the reference section lists only five post-1983 entries. Nevertheless the book is very relevant to the field of computational linguistics: it constitutes an archetype of the assumptions, strategies, and limitations faced by anyone attempting to implement a text-processing tool.

This relatively short book (188 pages, double-spaced) is divided into two parts. The first, comprising four chapters, overviews the model, which is then detailed in the second part. The first chapter introduces the basic assumptions and goals of the thesis; the research focuses on memory mechanisms, not inferencing or reasoning. The author states that the work is carried out in terms of automatic natural-language processing (NLP) and thus that he will avoid claims and suggestions about human language processing. Indeed, the title is somewhat misleading: the use of the word "memory" in the dissertation has little to do with human memory. In

essence, the thesis describes marker-passing algorithms used to select between possible candidates for disambiguation. The algorithms search for candidates in a database and choose between them according to the current context, which simply constrains memory retrieval. The system, called **Capture**, was designed not only for text processing but also to process collections of English paragraphs and produce an output incorporable into a conventional database.

The second chapter overviews the representational scheme and algorithms developed by the author. The knowledge base is constructed out of two types of assertions: **specializations** (IS-A declarations) and **correspondences**, which take the form *role C1 of owner D1 is a role-specialization of role C2 whose owner is D2*. These types of assertions can carry further information about the relationships between their arguments. This information is encoded as a list of flags given as an additional argument to the assertion. The author remarks that "the motivation for the choice of flags that were defined for the memory formalism is simply that these seem to be useful, in practice, for stating information at the level of this kind of formalism". Context is represented by a collection of context factors, each of which contributes activation to a particular set of memory entities (i.e., to the "objects" referred to in the assertions). There are seven major types of context factor, including recency, syntactic emphasis, deixis, and a priori subject area. These are essentially static rules that define how activation is managed for each memory entity involved during comprehension. The rules are applied for disambiguation and for defining the focus space; that is, the set of most "activated" memory entities. In the remainder of the chapter, Alshawi discusses his standard marker-passing model.

The third chapter addresses the problem of interpretation. The author tackles noun phrase (NP) reference interpretation, compound NPs, possessive NPs, *with*-PPs, and word-sense disambiguation. Conditionals, negation, and "phenomena going beyond memory mechanisms" (e.g., modality and metaphor) are not handled. The algorithms are simple to understand but often lack proper motivation. In the fourth chapter, Alshawi discusses related research as of 1983. In particular, the author acknowledges the strong influence of Fahlman's work on marker-passing, and of Grosz's notion of **global focus**.

The second part starts on page 76. In the rest of the book, Alshawi details the ideas of the first part (see summary table, p. 94) and elaborates on the Capture feature that creates a relational database as the result of text processing. The author concludes with a chapter on the complexity of techniques for efficient retrieval from a database, a topic too often ignored in NLP models.

I said above that this book is, in my opinion, an archetype of the NLP thesis in computational linguistics. The first appendix, which lists some of the 30 short texts processed by Capture, confirms this: the examples

are incredibly artificial! As with several other AI systems, one gets the distinct impression that the data is fitted to the software! Alshawi is looking for the algorithms that produce the optimal interpretation of a text. How can he justify these algorithms when he discards psychology and neurophysiology? What constitutes the optimal interpretation of a text? These questions are left unanswered. What remains is the implementation of a solution to some specific problems considered in a vacuum. The dissertation presents the advantages and limitations of a particular solution; several other solutions to these problems have been suggested. Each author compares his model to the others but the relative vacuity of it all persists. Suspiciously, all models share the same basic flaws (e.g., simplistic model of memory, and absence of mechanisms for the subsequent correction of an erroneous interpretation). By dissociating NLP, with its cartesian quest for optimal interpretation algorithms, from man, his language, his formidable ability to understand and to misunderstand and to not understand, computational linguists seem to have created the ultimate field of study, one where partial solutions are taken to intrinsically hold the promise of an eventual complete and correct solution to the problem of linguistic understanding. I reject the notion of the "optimal interpretation" of a text and, after reading Alshawi's book, I am left with the bitter taste of an interesting yet very artificial Lisp program.

*Jean-Pierre Corriveau* is a doctoral candidate in the Department of Computer Science, University of Toronto, working on a psychologically valid model of schematic memory for text understanding. He is also a member of the Exploratory Tools group of Bell-Northern Research, Ottawa. Corriveau's address is: 748 Parkdale Ave., Ottawa, Canada K1Y 1J9. Email: jpierre @ ai.toronto. edu.

---

**FROM TEXT TO SPEECH: THE MITALK SYSTEM**  
(CAMBRIDGE STUDIES IN SPEECH SCIENCE AND COMMUNICATION)

**Jonathan Allen; M. Sharon Hunnicutt; and Dennis Klatt, with Robert C. Armstrong and David Pisoni**

(Respectively: MIT; Royal Institute of Technology, Stockholm, Sweden; MIT; MIT; Indiana University)  
Cambridge, England: Cambridge University Press, 1987, xi + 216 pp.  
ISBN 0-521-30641-8, \$29.95 (hb)

*Reviewed by*  
*Matti Karjalainen*  
*Helsinki University of Technology*

Speech processing and computational linguistics have traditionally been disciplines separated by different approaches and methodologies. The continuous nature of speech signals is not easily interfaced to the discrete

units and symbolic processing of linguistic and conceptual levels. One of the successful contributions in combining language and speech aspects is MITalk, a speech-synthesis system developed at MIT to convert English text to speech signals. The book *From text to speech: The MITalk system* is a detailed documentation of the principles that have resulted from a research effort that has lasted more than two decades.

The MITalk project has influenced the theory and practice of speech synthesis to a large extent. Several commercial synthesizers draw their origin from MITalk and many research groups around the world have benefited from its underlying principles, even those involved in languages as different as Finnish—in the case of the reviewer. The MITalk synthesizer is regarded as a reference point in this field and this systematically written book on it is highly welcome.

The writers do not attempt to give a comprehensive review of different approaches and implementations available in speech synthesis. (For this, see Klatt 1987.) Instead, they build their own theoretical framework by a thorough description of the MITalk implementation principles. The book is divided into parts in the same way as the synthesis process of MITalk itself. The book, as well as the text-to-speech synthesis, starts from the analysis of the text input. Text pre-processing is followed by morphological analysis. The extensive morph lexicon contributes significantly to the high quality of the synthesis process. A phrase-level parser is used with rules for morphophonemics and stress modification. The application of ordered phonological rules to letter-to-sound conversion and lexical stress concludes the text-analysis part of the text-to-speech process.

The synthesis part of MITalk contains a phonological component, a prosodic component including a fundamental frequency generator, a phonetic component (all the time approaching a continuous-time representation)—and finally a formant synthesizer (called the Klatt synthesizer) to produce the continuous speech signal. The rest of the book discusses some pragmatic aspects like measures of intelligibility of the synthetic speech.

This book is primarily of interest to those who are active in speech research and synthesizer development. However, especially the first half of the book might draw remarkable attention from the point of view of computational linguistics. Most chapters are easy to read (a possible exception is the formant-synthesizer section). One weakness of the book is that it took too many years to publish after the most intensive phases of the research work. There are also too few ideas to be found from an artificial-intelligence perspective even though the rule-based processing in speech synthesis could be improved by modern knowledge-based methods. The main message of the book to computational linguistics is perhaps threefold. First, that a close relationship to speech processing exists; second, it should