

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

be known; and finally, that it should be kept alive in research since it exists in the real human-communication process itself.

#### REFERENCE

Klatt, Dennis, 1987. Review of text-to-speech conversion for English. *The Journal of the Acoustical Society of America* 82(3): 737–793.

*Matti Karjalainen* is a professor at the Helsinki University of Technology, Finland. His research activities include speech synthesis, analysis, recognition, and auditory modeling of speech perception. Karjalainen's address is Helsinki University of Technology, Acoustics Laboratory, Otakaari 5A, SF-02150 Espoo, Finland.

#### JAPANESE PHRASE STRUCTURE GRAMMAR: A UNIFICATION-BASED APPROACH (STUDIES IN NATURAL LANGUAGE AND LINGUISTIC THEORY)

Takao Gunji (Faculty of Language and Culture, Osaka University, Japan)

Dordrecht, Holland: D. Reidel, 1987, ix + 239 pp. ISBN 1-55608-020-4, \$64.50 (hb)/£49.50, Dfl 135.

*Reviewed by*  
*Pete Whitelock*  
*University of Edinburgh, Scotland*

Professor Gunji's book continues the tradition of explicitness in linguistic description so well demonstrated by Gazdar, Klein, Pullum, and Sag (1985) (GKPS). In theoretical orientation, too, there is a heavy debt to the latter. The details, however, show a certain trend towards the lexicalist position exemplified in theories such as head-driven phrase-structure grammar (Pollard and Sag 1987), and categorial unification grammar (Uszkoreit 1986, Karttunen 1986, Zeevat et al. 1986).

All these theories have been developed with a strong computational orientation. GKPS could be considered essential reading for anyone intending or attempting to develop a serious computational grammar for English, and Gunji's Japanese Phrase Structure Grammar (JPSG) stands in the same relation to Japanese. In fact, JPSG is used as the basis of the natural-language component of the Japanese Fifth Generation Computer project.

The book, a considerably revised version of the author's Ohio State M.A. thesis, presents a thorough and extensive analysis of the central syntactic phenomena of the Japanese language—control in complementation, reflexive and zero anaphora, unbounded dependencies, and word-order variation.

After a brief introduction, the apparatus to be used in the analysis is presented. Despite the book's title, the phrase-structure rule component is reduced to a single schematic rule,  $M \rightarrow D H$ . The burden of constraining the well-formedness of the local trees licensed by this

rule falls on several principles that govern the distribution of features in the categories M(other), D(aughter), and H(ead). These include the **head-feature principle** (HFP) and **foot-feature principle** (FFP), familiar from GKPS; an **adjunct-feature principle**, which requires the value of the ADJUNCT feature of the category D to unify with the category H; and a **SUBCAT feature principle** (SFP), which ensures the satisfaction of lexical complementation requirements, as in HPSG. SUBCAT is a feature that takes sets of case-marked categories as its value, and the SFP states that in complementation, the value of the mother's SUBCAT is that of the head minus the element that unified with the non-head daughter.

Gunji assumes for most of the book a semantics with  $\beta$ -reduction of  $\lambda$ -calculus expressions as the basic combination operation, and it is this that is explained in the preliminaries section. However, in the final chapter, on word order, he suggests that a semantic combination based on unification is more appropriate for handling a language with the word-order characteristics of Japanese.

Chapter 3, "Fundamental constructions", starts with a detailed argument for the existence of a verb phrase (VP) in Japanese, and by corollary for the importance of *subject*, since VP abbreviates the category  $V\{SUBCAT\{PP\{SBJ\}\}\}$ . The rest of the book crucially exploits the notion of VP. The remainder of Chapter 3 examines the productive processes of causativization, passivization, and benefactivization. These phenomena have been the focus of continuing research since Kuroda 1965, and have generally been treated by considering the causative, passive, and benefactive morphemes as verbs taking sentential complements in the deep structure, with obligatory *equi* or raising and predicate raising. In Gunji's analysis, these morphemes subcategorize for subject, (possibly dative) object, and a (crucially) unsaturated verbal constituent. The latter is either a VP (=  $V\{SUBCAT\{PP\{SBJ\}\}\}$ ), in causatives and intransitive (adversity) passives, or a TVP (=  $V\{SUBCAT\{PP\{SBJ\}, PP\{OBJ\}\}\}$ ) for the transitive passive.

Given the unification-based semantics of the final chapter, the object control is captured simply by identifying the semantics of the object and the VP or TVP's subject in the lexical entries for the bound morphs. Gunji explicitly declines to give a consistent account of case marking in simplex sentences and these constructions, thus failing to explain why the causee may not be marked with accusative case when the embedded verb is transitive (the latter fact being captured with a feature co-occurrence restriction.) Neither does he appear to be worried by the existence of a phrase boundary occurring internally to a word, as in:

Ken ga Naomi ni [vp hon wo yom] ase ta  
NOM DAT book ACC read CAUSE PAST  
(Ken made Naomi read the book.)