

# Book Reviews

## Cognitive Models of Speech Processing

Gerry T. M. Altmann (editor)  
(University of Sussex)

Cambridge, MA: The MIT Press  
(The ACL–MIT Press Series in Natural  
Language Processing), 1990, x + 540 pp.  
Hardbound, ISBN 0-262-01117-4, \$55.00

*Reviewed by*  
*Howard S. Kurtzman*  
*National Institute of Mental Health*

This volume is one of the most valuable publications in psycholinguistics of the last decade. It is a collection of 22 chapters, all written by leading contemporary investigators, presenting much of the best recent empirically-based research and speculation on spoken language comprehension by humans (mainly adults). The chapters divide into three general areas, all of primary current concern: word identification (occupying about two-thirds of the book), syntactic parsing, and the relation between intonational structure and syntactic structure. Literally every chapter develops interesting ideas in a clear and rigorous fashion.

The volume is particularly useful because, unlike most collections, the chapters engage one another substantively. Three chapters act as explicit commentaries on others, and elaborate independent points as well. Many of the other chapters also contain specific responses to the claims and approaches described in related chapters. Further, the editor's introduction comprehensively delineates the areas of agreement, disagreement, and complementarity among the chapters. As a whole, the book provides a coherent picture of the rich theoretical and methodological scene that comprises psycholinguistics today.

In the following, I summarize the background and major themes of each group of chapters.

### 1. Word Identification

Current theoretical and methodological approaches to auditory word identification are strongly inspired by the work of Marslen-Wilson (1987) and his colleagues. This work indicates that upon reception of the initial sounds of a word, all of the words in the perceiver's lexicon that share those initial sounds are activated in parallel. These words might vary in their degree of activation, depending upon how closely they fit the auditory stimulus. As further sounds are received, the activation levels are adjusted, until a point is reached (perhaps before the end of the word stimulus) at which a decision procedure determines that one word has significantly greater activation than any others and adopts it as the lexical percept. In addition, syntactic and semantic context can influence activation levels or the decision procedure.

Arising from this general conception are a number of specific issues addressed by the contributors (Richard Shillcock; Uli Frauenfelder and Guus Peeters; Dennis Norris;

Anne Cutler; Paul Luce, David Pisoni, and Steven Goldinger; William Marslen-Wilson; Jan Charles-Luce, Paul Luce, and Michael Cluff; Ellen Bard; Gerry Altmann; Jacques Mehler, Emmanuel Dupoux, and Juan Segui [two chapters]; Cynthia Connine; Arthur Samuel; Lorraine Tyler; Henry Thompson and Gerry Altmann; Jeffrey Elman):

1. What algorithm underlies the decision procedure for adoption of a particular word?
2. It is already known that the frequencies of occurrence of the word candidates influence the speed of word identification. What is the locus of this influence: initial activation levels or the decision procedure?
3. What influence, if any, does the size of the set of activated word candidates have upon identification processes?
4. Do strongly activated word candidates inhibit the activation levels of other candidates?
5. What is the role of the syllable, particularly the stressed syllable?
6. How are word boundaries in the continuous speech stream located?
7. What is the nature and degree of contextual influences on word identification? (This is the issue of modularity, various characterizations of which are considered by the authors.)

The chapters provide contrasting responses to each of these issues, with many chapters addressing more than one issue. Data are reported from psycholinguistic experiments on humans, cognitive simulations (which are really another form of psycholinguistic experimentation), and statistical analyses of dictionaries. Much of the research aims to develop precise connectionist models, with the major influences being McClelland and Elman's (1986) TRACE model and Jordan's (1986) recurrent-link dynamic networks. Furthermore, Elman's chapter, which is framed as a reply to Fodor and Pylyshyn (1988), demonstrates just how radical an alternative to traditional symbol-based processing connectionism can be (cf. related work by Elman (1990) and van Gelder (1990)).

## 2. Syntactic Parsing

Some theorists (e.g., Clifton and Frazier 1989) have claimed that only general structural principles and strategies are used by initial parsing processes, with application of lexical information delayed until a later stage of parsing. The chapter by Michael Tanenhaus, Susan Garnsey, and Julie Boland presents experimental evidence (including neurophysiological results) against this claim. Their work indicates that locating *wh*-gaps is immediately guided by the parser attempting to assign the thematic roles specified by the verb (see also Tanenhaus and Carlson 1989; Kurtzman, Crawford, and Nychis-Florence 1991). Their experiments also indicate that control relations (as in *Jane forced/promised Mary to go to the store*) are immediately computed on the basis of the verb.

Lyn Frazier's chapter develops a view of the parser as modularly organized. She argues for separate modules for: building constituent structure, assigning binding relations (e.g., in *wh*-constructions), assigning thematic roles, and assigning syntactic

constituents to discourse elements (cf. Weinberg 1987). Frazier claims that the first two of these are completely “informationally encapsulated,” which is one of Jerry Fodor’s (1983) criteria for a module. The other two, however, are only “pseudoencapsulated,” since they make constrained use of nonlinguistic information (but otherwise obey Fodor’s criteria). They serve as intermediaries between linguistic and general knowledge (similar to the claims of Tanenhaus and Carlson 1989).

The commentary chapter by Janet Dean Fodor analyzes and extends the proposals of Tanenhaus et al. and Frazier in terms of the broad issues that organize recent parsing research, including single versus parallel analysis and the role of semantic influences.

The claims put forward in these chapters are not specific to spoken language processing. Researchers today assume that most parsing processes operate on both spoken and written input, and in fact parsing experiments typically use only written stimuli (because controlled presentation is easier to maintain). However, spoken language does uniquely give rise to questions concerning the role of intonational information in parsing.

### 3. Intonation

In recent years the dominant linguistic approach to explaining intonation phenomena has been that represented by Selkirk (1984), who posits distinct levels of intonational structure, syntactic structure, and information structure (concerned with focus/presupposition). (See also Inkelas and Zec 1990.) Mark Steedman’s chapter argues that the alternative approach of combinatory categorial grammar (Steedman 1989) allows these three levels to be collapsed into one. This then opens up the possibility of a sentence comprehension system that uses syntactic, intonational, and semantic information simultaneously and efficiently in the word-by-word processing of the incoming sentence (see also Altmann and Steedman 1988). Aravind Joshi’s commentary chapter argues that tree-adjointing grammars can be developed that are weakly equivalent to combinatory categorial grammars and that also maintain a close relation between syntactic and intonational structure (although not the identity relation as proposed by Steedman).

Finally, the chapter by Mitchell Marcus and Donald Hindle offers a parser-based account of obligatory intonation boundaries (those boundaries whose absence would render the string ungrammatical or would require that a different syntactic structure be assigned to it). Marcus and Hindle assume, following work such as that of Tanenhaus and his colleagues, that the syntactic parser actively assigns a verb-specified thematic role to each constituent as it is identified. However, they further assume that an intonation boundary terminates thematic assignment for a particular verb. Marcus and Hindle argue, within the deterministic parsing framework of description theory (Marcus 1987), that obligatory boundaries occur at all and only the points that are necessary to prevent misassignment of thematic roles. (See related experimental work by Carroll and Slowiaczek 1987.)

### 4. Conclusion

This volume provides complete state-of-the-art coverage (as of 1990) of research in auditory word identification, including specific methodologies and data. Its coverage of parsing and intonation research is not as broad or detailed, but the chapters do touch upon all the chief issues and present important ideas that will be influential in future work.

Although the volume is primarily a contribution to the professional psycholinguistics literature, most of it is accessible to any motivated reader who is generally acquainted with the issues underlying recent investigation of speech and language processing. Many of the chapters are admirable for their careful organization of complex material. The editor's introduction is an accurate and useful overview of the volume, although it provides not so much an orientation or grounding as a high-level summary that makes sense only after the chapters have been skimmed for their main points.

As this volume demonstrates, psycholinguistics has grown enormously in both range and sophistication since the days of Fodor, Bever, and Garrett (1974). Not only have new empirical methods been developed, but considerable consensus has been attained on the conceptual frameworks within which specific theoretical claims are formulated and tested. (Nonetheless, one broad conceptual issue touched upon in this volume for which there is not yet consensus is whether connectionist representations can adequately capture the linguistic knowledge that has traditionally been expressed in the form of explicit symbolic rules.) Further, there has been a significant convergence on particular classes of processing models. The outlook for a cognitive theory of language has never been brighter.

### References

- Altman, G., and Steedman, M. (1988). "Interaction with context during human sentence processing." *Cognition*, 30, 191–238.
- Carroll, P. J.; and Slowiaczek, M. L. (1987). "Modes and modules: Multiple paths to the language processor." In *Modularity in Knowledge Representation and Natural-Language Understanding*, edited by J. L. Garfield. Cambridge, MA: The MIT Press, 221–247.
- Clifton, C.; and Frazier, L. (1989). "Comprehending sentences with long-distance dependencies." In *Linguistic Structure in Language Processing*, edited by G. N. Carlson and M. K. Tanenhaus. Dordrecht: Kluwer, 273–317.
- Elman, J. L. (1990). "Finding structure in time." *Cognitive Science*, 14, 179–211.
- Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: The MIT Press.
- Fodor, J. A.; Bever, T. G.; and Garrett, M. F. (1974). *The Psychology of Language*. New York: McGraw-Hill.
- Fodor, J. A.; and Pylyshyn, Z. W. (1988). "Connectionism and cognitive architecture: A critical analysis." *Cognition*, 28, 3–71.
- Inkelas, S.; and Zec, D. (eds.) (1990). *The Phonology–Syntax Connection*. Chicago: University of Chicago Press.
- Jordan, M. I. (1986). "Serial order: A parallel distributed processing approach." Technical Report 8604, Institute for Cognitive Science, University of California, San Diego.
- Kurtzman, H. S.; Crawford, L. F.; and Nychis-Florence, C. (1991). "Locating *wh*-traces." In *Principle-Based Parsing*, edited by R. C. Berwick, S. P. Abney, and C. Tenny. Dordrecht: Kluwer, 347–382.
- McClelland, J. L.; and Elman, J. L. (1986). "The TRACE model of speech perception." *Cognitive Psychology*, 18, 1–86.
- Marcus, M. (1987). "Deterministic parsing and description theory." In *Linguistic Theory and Computer Applications*, edited by P. Whitelock. London: Academic Press.
- Marslen-Wilson, W. D. (1987). "Functional parallelism in spoken word-recognition." *Cognition*, 25, 71–102.
- Selkirk, E. O. (1984). *Phonology and Syntax*. Cambridge, MA: The MIT Press.
- Steedman, M. (1989). "Coordination and constituency in a combinatory grammar." In *Alternative Conceptions of Phrase Structure*, edited by M. Baltin and A. Kroch. Chicago: University of Chicago Press.
- Tanenhaus, M. K.; and Carlson, G. N. (1989). "Lexical structure and language comprehension." In *Lexical Representation and Process*, edited by W. D. Marslen-Wilson. Cambridge, MA: The MIT Press.
- Van Gelder, T. (1990). "Compositionality: A connectionist variation on a classical theme." *Cognitive Science*, 14, 355–384.
- Weinberg, A. (1987). "Modularity in the syntactic parser." In *Modularity in Knowledge Representation and Natural-Language Understanding*, edited by J. L. Garfield. Cambridge, MA: The MIT Press, 259–276.

*Howard S. Kurtzman* received his Ph.D. in cognitive psychology from MIT and has taught at Cornell University. He is now at the National Institute of Mental Health (Basic Brain and Behavioral Sciences Division), 5600 Fishers Lane, Rockville, MD 20857; e-mail: [hskx@vax5.cit.cornell.edu](mailto:hskx@vax5.cit.cornell.edu)