LEXICAL AMBIGUITY AS A TOUCHSTONE FOR THEORIES OF LANGUAGE ANALYSIS

Lawrence Birnbaum

Yale University
Department of Computer Science
New Haven, Connecticut

Abstract

This paper assesses several broad approaches to language analysis with respect to the problem of lexical ambiguity. The impact of the problem on both syntactic and semantic analysis is discussed, and several common methods for disambiguation, including the use of selectional restrictions and scriptal lexicons, are analyzed. Their shortcomings illustrate the need for complex inference to resolve ambiguity, which forms one of the key functional arguments in favor of integrating language analysis with memory and inference. However, it has proven surprisingly difficult to realize such an integrated approach in practice: An assessment of lexical disambiguation within some recent models which attempt to do so reveals that they rely largely on the traditional techniques of selectional restrictions and scriptal lexicons, with all their drawbacks. The difficulty is shown to stem primarily from the theories of memory and inferential processing utilized. The implications for recent approaches to language analysis based on connectionist mechanisms are explored. Finally, the requirements imposed by lexical disambiguation on theories of memory and inferential processing are discussed.

Introduction

The problem of natural language analysis, or "parsing," has been approached in many different ways and from the perspective of many different theoretical traditions. Because these theoretical traditions often differ quite radically in their basic assumptions about the goals of language analysis and the methods that ought to be employed, it can be quite difficult to compare the different approaches. But regardless of these differences, there are certain characteristics of the input that must be dealt with. Natural language is elliptic, ambiguous, and vague, to name just three of these problematic features Any language analyzer must contend with some or all of these problems. This suggests that one good way to try and make sense out of the variety of approaches is to examine their various strengths and weaknesses with regard to such characteristics.

In this paper, I propose to evaluate several broad approaches to parsing with respect to one of the most basic of these problematic characteristics, lexical ambiguity. Lexical ambiguity is one of the chief sources of ambiguity in language, so the problem is undeniably important. It is, further, widely recognized to be a far more pervasive phenomenon than it intuitively seems to be. Because people are not consciously aware of most of the ambiguities in what they read or hear, the fact that most of what they read or hear *is* ambiguous is not immediately apparent. However, a glance at any ordinary dictionary should make it plain that lexical ambiguity is extremely common.

Lexical ambiguity is, finally, a problem the importance of which has long been appreciated. It was one of the rocks on which the early work in machine translation foundered. Bar-Hillel (1960), in his critique of that work, showed that determining the correct sense of an ambiguous word depends, in general, on plausible inferences from extremely complex features of the context in conjunction with arbitrary facts about the world. He gave as an example the problem of choosing the correct meaning of the word "pen" in the sentence "The box is in the pen." In this sentence, the pen in question is probably an enclosure, such as a play-pen, rather than a writing implement. Bar-Hillel argued that in order to determine this, a language analyzer would need access to knowledge of the functions and relative sizes of these two different kinds of objects, as well as some means of using that knowledge to determine the plausibility of the various possible interpretations of the sentence

Of course, lexical ambiguity is not just a problem for *semantic* analysis. It is also one of the chief causes of structural ambiguity, and it is, therefore, an issue with which syntactic analyzers must contend as well. This aspect of the problem has also long been appreciated. In the well-known example "Time flies like an arrow," (Kuno, 1965), much of the structural ambiguity of the sentence stems from the part-of-speech ambiguity of the words "time," "flies," and "like," which in turn reflects their semantic ambiguity.

In sum, the problem of lexical ambiguity can indeed serve as a touchstone by which theories of language analysis can be assessed. The problem is basic and pervasive. The issues implicated in its solution, and the problems to which it gives rise, have long been appreciated. It arises regardless of whether one is trying to construct a syntax-based parser or a semantics-based one. Despite its importance, however, surprisingly little progress has been made on the problem. In this paper, I will attempt to provide a critical survey of what has been accomplished. No new solutions will be presented. However, the critique will reveal some of the requirements

for a solution, and some of the consequences for the understanding process as a whole will be explored

Lexical ambiguity and syntactic analysis

In syntactic analysis, the problem of lexical ambiguity is not the problem of choosing the correct sense of a word, but simply the correct part of speech. However, as the last example demonstrated, these problems are not. unrelated. Word-sense ambiguity very often entails part-of-speech ambiguity as well. Thus, correctly disambiguating the part of speech of a word will in general depend on complex semantic and pragmatic processing Syntactic analyzers cannot, therefore, be expected to solve by themselves the problem of lexical ambiguity, even just part-of-speech ambiguity. It is not unreasonable, however, to expect that they might contribute to its solution.

The chief approach to resolving syntactic ambiguity, lexical or otherwise, is simply to try each alternative, while being prepared to back up in case it should prove mistaken. This is the approach taken in ATN parsers and descendant models (see, e.g., Thome, Bratlcy, and Dewar 1968. Bobrow and Fraser, 1969; Woods, 1970, Fereira and Warren, 1980) When such a parser encounters an ambiguous word, it simply tries each possible choice for that words part of speech which will enable a transition, and which therefore offers the possibility of successfully parsing the input sentence according to the grammar utilized. If the choice does not lead to a successful syntactic analysis, then it will be discarded when the parser backs up. (By performing an incremental semantic analysis on structures proposed by the syntactic analyzer, it is possible to rule out choices on semantic grounds as well; see. eg., Bobrow and Webber, 1980.) This process will be repeated for a given word each time the parser encounters it when driving forward in the network All and only the choices that lead to successful analyses will be output with those analyses. Further disambiguation is the responsibility of the semantic and pragmatic components of the understanding process

More recently, however, Marcus (1980) has criticized this approach to resolving ambiguities, and has argued instead that syntactic analysis can normally be accomplished without resorting to unlimited back-up. In particular, he claims that syntactic structural ambiguities must and can be resolved with limited look-ahead and highly restricted use of semantic information. Since much of the structural ambiguity in language arises as a result of lexical ambiguity, lexical ambiguity is clearly one of the crucial issues which must be faced in making such a claim Nevertheless, Marcus's theory barely addresses the problem: with only one or two exceptions, words are taken to be syntactically unambiguous in his work. At the very least, this failure to confront the issue makes it difficult to evaluate the status of the theory.

In fact, the one or two cases of lexical ambiguity which Marcus does attempt to resolve within the framework of his theory simply serve to show how-profound the impact of the problem actually is. For example, in order to disambiguate whether the word "have" is used as an auxiliary or a main verb, Marcus introduces a diagnostic rule which is arguably the most complex in his entire grammar. Nevertheless, as Marcus himself points out, the rule fails on many obvious examples. How well such rules would work in the context of many *other* ambiguous words is highly questionable. Indeed, Milne's (1982) attempt to address lexical ambiguity within the framework of Marcus's theory led to a substantially greater reliance on semantics. One need not agree with the details of his proposals to find this result suggestive.

Lexical ambiguity and semantic analysis

We have seen that syntactic analyzers, alone, cannot be expected to do very much about lexical ambiguity It is, after all, primarily a question of word-sense ambiguity rather than just part-of-speech ambiguity, and so primarily a semantic problem rather than a syntactic one. Quite naturally, therefore, it is an issue which has received far more attention in semantic analyzers than in syntactic ones. At first glance, there seem to be a variety of different semantic approaches to the problem. In fact, however, most approaches turn out to share only one or two fundamental mechanisms.

The major semantic approach to the solution of lexical ambiguity involves the use of *selectional restrictions* (Katz and Fodor. 1963) These are semantic requirements associated with the structures representing the meanings of words or phrases, which must be met by another semantic structure before the two can be combined. For example, an action like eating might require that its actor be animate In general, selectlonal restrictions are one-place predicates that test for the presence or absence of some semantic feature, or some boolean function of such predicates

The use of selectional restrictions in disambiguation is, in principle at least, quite straight forward. One simply chooses the sense (or senses) of a word that select ional restrictions will allow to combine with other semantic structures in the sentence, either because it meets the requirements of those other structures, or because they meet its own requirements. To paraphrase a simple example from Katz and Fodor (1963), consider the word "ball." This can mean, among other things, either a fancy party with dancing, or a round object used as a toy. In the sentence "John hit the ball," the use of selectional restrictions would result in choosing the round object sense of "ball," since the action of hitting can be applied to a physical object but not to a social gathering.

A variety of different methods have been developed for applying selectional restrictions in the resolution of lexical ambiguity; I will briefly mention just a few of them here. Winograd (1972) proposed that they be used by semantic interpretation specialists associated with functional syntactic constituents such as noun groups and clauses. Riesbeck (1975) proposed encoding selectional restrictions in the tests of the lexically indexed

productions that represent, in his theory, the different meanings of a word. Rieger and Small's (1979) theory of word experts and Hirst and Charniak's (1982) theory of Polaroid words are based on more sophisticated versions of this idea. Wilks (1976) has proposed that selectional restrictions should not be absolute requirements, but simply *preferences.* In his model, one picks the sense of each word that maximizes the total number of preferences satisfied in a given sentence.

The other major approach to handling lexical ambiguity involves the use of a scriptal *lexicon* (Schank and Abelson, 11)77, Cullingford, 1978; Riesbcck and Schank, 1978, Charniak 1981) This idea is based on the observation that many words have special meanings in particular contexts. Thus, in a sense, each script or frame used in understanding a text should have an associated lexicon in which words are assigned their frame-specific meaning. For example, the frame for a baseball game would have an associated lexicon in which the word "home" would be defined as the plate in the ground over which batters stand, and which a player must touch to score a run. By itself, this idea is not very useful for disambiguation, except insofar as it keeps frame-specific meanings out of consideration unless the relevant frame is "active." The crucial simplifying assumption which is usually made, therefore, is that if a given frame is "active," all words in its scriptal lexicon can be presumed to have their frame-specific meaning

Although both selectional restrictions and scriptal lexicons are very useful up to a point, especially in domain-limited applications, it should be clear that they have severe limitations. The simplifying assumption which underlies the scriptal lexicon approach, that words will not be used in other than their frame-specific sense, is clearly not true. For example, consider the following sentence in the context of a story about a baseball game: "The game was so lopsided that Fred got bored and walked home after the seventh inning." Here, the home in question is probably Freds residence, not home plate.

The use of selectional restrictions has similar limitations. Consider the following variant of Bar-Hillel's example: "The pen is in the box along with assembly instructions." Here, the pen in question is probably a play-pen, and almost certainly not a writing implement. Determining this requires recognizing that the assembly instructions are probably for the assembly of the pen, and knowing that play-pens often require assembly by the consumer after purchase, whereas writing implements do not Using this knowledge in turn requires inferring that since the pen is in a box with assembly instructions, it has probably just been purchased by the consumer. The point here is that these are simply not the kinds of rules that can be represented and employed as selectional restrictions, except at the risk of precluding the correct analysis of other examples. We cannot, for example, just invent a feature "objects that can be assembled" as a selectional restriction on the object of "assemble," and which would be a property of play-pens but not writing implements. Writing pens certainly *are* assembled, in factories, and they may even be assembled by the consumer, as in "John assembled the pen after cleaning it and putting in a new cartridge "

Lexical ambiguity and Integrated analysis

The above discussion makes it. clear that what must be brought to bear on the problem of lexical ambiguity are the general inference and memory processes used in understanding. Thus, lexical ambiguity is one of the key problems which motivates an integrated approach to language analysis, ONE- in which inference and memory processing play an important role in the analysis process itself (Schank, Lebowitz, and Birnbaum, 1980, Schank and Birnbaum, 1984*).* Although it plays a key role in motivating this approach, however, and would therefore seem crucial to theories of integrated analysis, surprisingly little attention has been devoted to it

For example, consider the approach taken in the model proposed by Dyer (1983), which explicitly aims to be a model in winch memory and inference are intimately entwined in the language analysis process Despite this intent, the discussion of lexical disambiguation in the model is limited to the use of selectional restrictions and scriptal lexicons Both are implemented as the tests of lexically indexed productions, in a manner similar to Riesbcck (1975) We can best see how this works by looking at a representative example For instance, here is the procedure which disambiguates the phrase "run into," slightly paraphrased for readability:

```
If the actor is a VEHICLE or
   the SCENARIO is TRANSITIONAL
        with a VEHICLE instrument,
Then interpret "run into" as VEHICLE-ACCIDENT.
Else If the object is
        a HUMAN who has an INTERPERSONAL
          RELATIONSHIP with the actor,
     Then interpret "run into" as
          RENEW-INTERPERSONAL-RELATIONSHIP
```

Lets analyze how this is intended to work The test for whether or not the actor is a vehicle is simply a selectional restriction. The test for whether "the SCENARIO is TRANSITIONAL with a VEHICLE instrument" is perhaps more puzzling However, its purpose would seem to be to handle examples such as "While I was driving home, I ran into a parked car,"[1] in which the actor of "run into" is not a vehicle, but the proper interpretation is nevertheless vehicle accident. In effect, this is an implementation of the scriptal lexicon idea: if the vehicle travel frame is active, then "run into" means vehicle accident. Both of these rules are subject to the limitations described in the last section. For example, this use of the scriptal lexicon approach would fail on the following text:

> While I was driving home, I remembered I needed some milk. I ran into a Seven-Eleven and picked up a half-gallon.

Finally, let's consider the test for a human who has some interpersonal relationship with the actor. Here, the model begins to employ knowledge beyond simple seleetional restrictions, which are technically just one-place predicates The problem is, it still *employs* thus knowledge exactly as if it were just a seleetional restriction Although the presence of such a relationship (or, in fact, any semantic feature) is indeed the sort of knowledge that may be *relevant* in determining the correct meaning of a word, its use as a sufficient condition in a non-inferential, lexically-indexed rule of this variety is entirely misplaced. The point LS that such knowledge must be represented and indexed in a way that makes *it* available for use by the general inferential capabilities of the understander

To be more specific about what is required, consider *how* the fact that two people have an interpersonal relationship might be relevant to determining the appropriate interpretation of "run into." If two people who knew each other happened to have a fortuitous encounter, then social rules such as politeness, and personal goals stemming perhaps from friendship, might *cause* them to pursue their interpersonal relationship at that juncture. They might, for example, engage in conversation, go to a bar, or arrange a subsequent meeting Knowledge of this causal relationship would enable an understander to explain why people who knew each other would exhibit such behavior, and thus enable the understander to construct a causally coherent representation of a textual fragment describing such an episode. It is the attempt to construct such a causally coherent representation that determines the proper interpretation of "run into" A particular interpretation, such as "social encounter," is preferred to the extent that it promotes such coherence

But the rule cited above does not explicitly represent such causal knowledge, nor does its choice of an interpretation for "run into" depend on the attempt to infer a causally coherent representation. Instead, the inference process is "short circuited" by directly linking some (but not all) of the relevant features with some (but not all) of the possible interpretations Such a rule simply cannot work in general. Consider, for example, the following text:

> John was racing down the street trying to catch a bus All of a sudden, his neighbor Fred stepped out of a doorway into his path. John ran into Fred and knocked him down. Fortunately, he wasn't hurt.

What both this example and the previous one demonstrate, to repeat, is that the proper interpretation of "run into" should be determined on the basis of the attempt by memory and inference to construct a causally coherent representation of the text as a whole — which is, after all, one of the chief functions of memory and inference in understanding. In a language analyzer which is truly integrated with memory and inference, it must be on the basis of these sorts of inferential considerations that language analysis problems, such as lexical ambiguity, are resolved Instead, in Dyers model we find that such inferential processing occurs *after* a word has already been disambiguated by means of seleetional restrictions and scriptal lexicons.

The model of integrated partial parsing proposed by Schank, Lebowitz, and Birnbaum (198G) and substantially extended by Lebowitz (1980) also depends, primarily, on the scriptal lexicon approach. In fact, most words are simply unambiguous as far as the model is concerned, since it presumes that input stories will involve only a single domain (terrorist incidents). To some extent, however, this model does make more serious use of memory and inference in disambiguation as well. In order to construct coherent representations of input stories, the model employs a version of script application (Schank and Abelson, 1977; Cullingford, 1978), in which an action or state is interpreted by matching a scriptal expectation The model can then use these expectations to disambiguate a word by choosing the meaning that satisfies one of them (This method was originally proposed by Riesbeck and Schank, 1978)

This method is clearly a step in the right direction. It is, however, subject to severe limitations, because it assumes, first, that if a script is active, then an ambiguous word *must* have the meaning that matches an expectation from that script, and second, that only one meaning will match an expectation But consider what would happen if more than one script were active, or if the scripts were larger and more detailed, or if expectations from sources other than scripts were utilized. U'nder these conditions, it seems quite likely that more than one meaning of an ambiguous word would match an expectation, or to put this another way, that more than one interpretation could be coherently interpreted within the context Thus, this method for using scriptal expectations will not work in many situations; it will either fail to disambiguate, or else simply choose in a way which guarantees a high probability of error. The method can only be employed reliably when only one script is active, and when only one sense of the word matches an expectation from that script As a result, this use of scriptal expectations is virtually equivalent, in the power and scope of its disambiguation capabilities, to the use of a scriptal lexicon For all practical purposes, one might just as well stipulate that the given word will have a given meaning if the given script is active.

Conclusions

How can the the use of scriptal expectations, or more generally of contextual expectations from varying sources, be extended to handle those cases in which more than one meaning of an ambiguous word might seem at first to fit the context? Several factors must be taken into account beyond the mere occurrence of a match between a potential word meaning and an expectation. First, which expectations are more important, or more likely to be satisfied at this point in the text? To put this in more general terms, which of the explanations for the different possible interpretations is more plausible or more salient? Second, does the text supply any additional clues? For example, a candidate semantic structure may be the right

sort of action to satisfy an expectation, but may nevertheless be inappropriate because its potential actor, as specified in the text, does not match the binding already assigned to the actor in that expectation. The use of such information is essential to exploit the full potential of memory and inference in lexical disambiguation.

In fact, this requirement poses the greatest challenge to recent models of language analysis employing connectlonist mechanisms (see, e.g., Small, Cottrell, and Shastri, 1982; Cottrell, 1984; Waltz and Pollack, 1984). The manipulation of variables and variable bindings is a difficult issue in the connectionist framework (J. Feldman, personal communication), and as currently formulated these models do not seem capable of utilizing such information in disambiguation. Thus, their use of contextual information in disambiguation seems subject to the limitations described at the end of the last section. Whether the clever manipulation of parameters such as weights and activation levels can overcome these limitations remains to be seen. One possible solution is to use connect lonist methods simply to suggest potential inference chains, and employ more traditional inference mechanisms, capable of manipulating variable bindings, to check over the suggestions (Charniak, 1983). Another possibility, requiring a more radical change in the connectionist framework, is to allow variable bindings to be passed between the units in a memory network (Riesbeck and Martin, 1985).

More broadly, however, the apparent difficulties in applying memory and inference to lexical disambiguation reflect not so much on the state of theories of language analysis as on theories of memory and inference. Here, the lesson of lexical ambiguity is that the knowledge needed to draw inferential connections in understanding cannot be packaged in isolated rules that commit the understander to certain inferences irrespective of what other rules may propose it is true that one possible interpretation of someone "running into" another person is as a fortuitous encounter leading to a social interaction. It is also true that one explanation for why two people would care to engage in such an interaction would be if they already knew each other. Thus, the knowledge that two people knew each other would provide support for interpreting "run into" as a social encounter, since such an interpretation would enable the understander to explain certain aspects of the situation. But, as we have seen, the decision that this is the *correct* interpretation cannot be made without considering the need to explain *other* aspects of the situation, aspects which may have nothing to do with social interactions and to which rules explaining such interactions cannot be expected to attend.

This last point bears particular attention. No single explanatory inference rule can be expected to attend to all the aspects of a situation which might affect the truth or relevance of the explanation it offers, and hence the validity of the interpretation it prefers for some vague or ambiguous linguistic element. Thus, determining which explanations to accept, and hence which interpretations to prefer, cannot be left to the inference rules themselves. Rather, there must be a more general inferential mechanism that determines which explanations to accept, taking into account the need to explain diverse aspects of a situation, and the evidence of diverse rules.

Probably the most ambitious attempt in this direction has been McDermott's (1974) model, which is capable of considering several potential explanations for a situation in parallel as it unfolds and choosing among them when evidence is available, as well as patching or replacing explanations that prove erroneous. Granger (1980) and ORorke (1983) propose models with this last capability as well, and Granger, Eiselt, and Holbrook (1984) have proposed a model of language understanding (including lexical disambiguation) which makes use of such techniques The most salient feature of these models is that they *explicitly* employ criteria, however crude, for deciding whether an explanation is adequate, when one explanation is preferable to another, and when an explanation has gone awry. For example, McDermott's criteria are, basically, coherence — an explanation must fit the facts — and parsimony -- an explanation with fewer unjustified assumptions is preferred. The use of such criteria would seem to be a crucial aspect of any inferential mechanism capable of fulfilling the requirements set out above, and thus capable of resolving lexical ambiguity in a general manner.

Acknowledgments: I thank Beth Adelson, Gregg Collins, and Alex Kass for useful discussions and for comments on an earlier draft of this paper. This work was supported in part by the Defense Advanced Research Projects Agency, monitored by the Office of Naval Research under contract N00014-85-K-0108

References

Bar-Hillel. Y. 1960. The present status of automatic translation of languages. In F. Alt, ed., *Advances in Computers* /, Academic Press, New York, pp. 91-163

Bobrow, D., and Fraser, B. 1969. An augmented state transition network analysis procedure. *Proceedings of the First IJCAI,* Washington, DC, pp. 557-567.

Bobrow, R , and Webber, B. 1980. Knowledge representation for syntactic/semantic processing. *Proceedings of the 1980 AAA1 Conference,* Stanford, CA, pp 316-323.

Charniak, E 1981. Six topics in search of a parser: An overview of AI language research. *Proceedings of the Seventh IJCAI,* Vancouver, B.C., pp. 1079-1087

Charniak, E. 1983. Passing markers: A theory of contextual influence in language comprehension. *Cognitive Science,* vol. 7, pp. 171-190.

Cottrell, G. 1984. A model of lexical access of ambiguous words. *Proceedings of the 1984 AAAI Conference,* Austin, TX, pp. 61-67.

Cullingford, R 1978. Script application: Computer understanding of newspaper stories. Research report no. 116, Yale University, Dept. of Computer Science, New Haven, CT

Dyer, M 1983. *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension* MIT Press, Cambridge, MA.

Granger, R. 1980. When expectation fails: Towards a self-correcting inference system. *Proceedings of the 1980 AAAI Conference,* Stanford, CA, pp. 301-305.

Granger, R , Eiselt, K., and Holbrook, J. 1984. The parallel organization of lexical, syntactic, and pragmatic inference processes. *Proceedings of the First Annual Workshop on Theoretical Issues in Conceptual Information Processing,* Atlanta, GA, pp. 97-106.

Hirst, G., and Charniak, *E* 1982. Word sense and case slot disambiguation. *Proceedings of the 1982 AAAI Conference,* Pittsburgh, PA, pp. 95-98.

Katz, J., and Fodor, J. 1963. The structure of a semantic theory. *Language,* vol. 39, pp. 170-210.

Kuno, S. 1965. The predictive analyzer and a path elimination technique. *Communications of the ACM,* vol. 8, pp. 453-462

Lebowitz, M 1980. Generalization and memory in an integrated understanding system. Research report no. 186, Yale University, Dept. of Computer Science, New Haven, CT.

Marcus, M. 1980. *A Theory of Syntactic Recognition for Natural Language.* MIT Press, Cambridge, MA.

McDermott, D. 1974. Assimilation of new information by a natural language-understanding system. Technical report no. 291, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Cambridge, MA.

Milne, R. 1982. Predicting garden path sentences. *Cognitive Science,* vol. 6, pp. 349-373.

(VRorke, P 1983. Reasons for beliefs in understanding: Applications of non-monotonic dependencies to story processing. *Proceedings of the 1983 AAAI Conference,* Washington, DC, pp 306-309.

Pereira, F., and Warren, D 1980. Definite clause grammars for language analysis - A survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence,* vol. 13, pp. 231-278.

Rieger, C, and Small, S. 1979. Word expert parsing. *Proceedings of the Sixth IJCAJ,* Tokyo, pp. 723-728.

Riesbeck, C. 1975. Conceptual analysis In R, Schank, ed., *Conceptual Information Processing,* North-Holland, Amsterdam, pp. 83-156.

Riesbeck, C, and Schank, R. 1978, Comprehension by computer: Expectation-based analysis of sentences in context. In W. Levelt and G. Flores d'Arcais, eds., *Studies in the Perception of language,* John Wiley, Chichester, England, pp. 247-293.

Riesbeck, C, and Martin, C. 1985. Direct memory access parsing. Research report no. 354, Yale University, Dept. of Computer Science, New Haven, CT.

Schank, R , and Abelson, R. 1977. *Scripts, Plans, Goals, and Understanding.* Lawrence Erlbaum, Hillsdale, NJ.

Schank, R., and Birnhaum, L. 1984. Memory, meaning, and syntax. In T Bever, J. Carroll, and L. Miller, eds., *Talking Minds: The Study of Language in the Cognitive Sciences,* MIT Press, Cambridge, MA, pp. 209-251.

Schank, R , Lebowitz, M., and Birnbaum, L. 1980. An integrated understander. *American Journal of Computational Linguistics,* vol. 6, pp. 13-30.

Small, S., Cottrell, G., and Shastri, L. 1982. Toward connectionist parsing. *Proceedings of the 1982 AAAI Conference,* Pittsburgh, PA, pp. 247-250.

Thome, J., Firatley, P , and Dewar, H. 1968. The syntactic analysis of English by machine. In D. Michie, ed., *Machine Intelligence 3,* American Elsevier, New York, pp. 281-309.

Waltz, D., and Pollack, J. 1984. Phenomenologically plausible parsing *Proceedings of the 1984 AAAI Conference,* Austin, TX, pp. 335-339.

Wilks, Y. 1976 Parsing English II In E. Charniak and Y. Wilks, eds., *Computational Semantics,* North-Holland, Amsterdam, pp. 155-184.

Winograd, T 1972 *Understanding Natural Language.* Academic Press, New York.

Woods, W. 1970. Transition network grammars for natural language analysis. *Communications of the ACM,* vol. 13, pp. 591-606.