

5. Sublanguages

Richard Kittredge, Chairperson

Université de Montreal
Montreal, Quebec PQH3C3J7

Panelists

Joan Bachenko, Naval Research Laboratory
Ralph Grishman, New York University
Donald E. Walker, SRI International
Ralph Weischedel, University of Delaware

5.1 Why Are Sublanguages Important for Applied Computational Linguistics?

Four of the five panels at this workshop are assessing the perspectives in applied computational linguistics for four important problem areas: natural-language interfaces, machine translation, text generation, and concept extraction. For each of these areas, it is assumed that any applied system will be oriented toward the particular variety of natural language associated with a single knowledge domain. This follows from the now widely accepted fact that such systems require rather tight, primarily semantic, constraints to obtain a correct analysis, and that such constraints can at present be stated only for sublanguages, not for the language as a whole. Although a practical system may well have components that are designed to accommodate the whole language, it must also anticipate the particular syntactic, lexical, semantic, and discourse properties of the sublanguage in which it will operate.

Research into the linguistic structure of weather reports, medical records, and aircraft maintenance manuals has led to specialized grammars for the sublanguages of these domains. Central to each sublanguage grammar is a statement of the functionally similar word classes and the co-occurrence restrictions among these classes. When a parser, generator, or translation system incorporates such a precise linguistic description, it becomes not only more efficient but also capable of discriminating between sentences (and texts) that are appropriate to the domain and those that are grammatical but inappropriate. In addition, the word classes used in the grammar, and the hierarchies relating these classes, are an important part of the knowledge structure for the domain.

5.2 How Do Sublanguages Arise?

When natural language is used in a sufficiently restricted setting, we may be justified in calling the resultant forms a sublanguage. Although there is no generally accepted definition of this term. Several

factors are usually present when the subset of a natural language is restricted enough for efficient semantic processing.

- ▶ Restricted domain of reference. The set of objects and relations to which the linguistic expressions refer is relatively small.
- ▶ Restricted purpose and orientation. The relationships among the participants in the linguistic exchange are of a particular type and the purpose of the exchange is oriented towards certain goals.
- ▶ Restricted mode of communication. Communication may be spoken or written, but there are constraints on the form of expression, which may include "bandwidth" limitations. Compressed (or telegraphic) language forms may reflect the time and space constraints of certain communication modes.
- ▶ Community of participants sharing specialized knowledge. The best canonical examples of sublanguages are those for which there exists an identifiable community of users who share specialized knowledge and who communicate under restrictions of domain, purpose, and mode by using the sublanguage. These participants enforce the special patterns of usage and ensure the coherence and completeness of the sublanguage as a linguistic system.

5.3 Constraints and Extensions in the Grammar of a Sublanguage

A typical sublanguage makes use of only a part of the language's lexical, morphological, syntactic, semantic, and discourse structures. These restrictions on its grammar, once detected and encoded in the form of rules, can be exploited during automatic processing by greatly reducing the number of possibilities to be considered. A sublanguage may also exhibit structures (and, hence, rules) that are not normally regarded as part of the standard language. In the most general case, then, a sublanguage grammar intersects, but is not contained in, the grammar of the general or standard language from which it derives.

Some of the typical constraints and extensions found in each component of a sublanguage grammar are given below, along with reference to recognized techniques for describing the constraints and for identifying them in a corpus of texts, when appropriate. In

addition, we mention a number of mechanisms for capturing these constraints for the purposes of computer processing.

5.3.1 Lexical and Morphological Characteristics

The most obvious feature of a sublanguage is its specialized lexicon. Not only is the set of word forms (and their possible meanings) highly restricted, but the productive word-formation rules may be of a particular kind, sometimes unique to the sublanguage or to a family of related sublanguages. Texts in medicine and pharmacology, for example, may contain a rich variety of names for diseases and drugs, which are constructed using characteristic affixes. Military sublanguages make frequent use of acronyms which obey describable rules of noun phrase formation in the grammar. Many sublanguages employ symbolic expressions (e.g., mathematics) or abbreviations which can be shown to have their own morphological characteristics.

Techniques for identifying the special morphology of sublanguage terms are readily available from linguistics. In cases where the lexicon is large, the designer of a computational system may find it profitable to include word-formation rules in a special processing phase.

5.3.2 Syntactic Characteristics

Early work on restricted language has shown that the syntactic description of a naturally occurring sublanguage may differ significantly from that of an unrestricted language. In the highly constrained style of weather bulletins, there is little resemblance between the syntactic structure of telegraphic forecasts and that of general language. The syntactic rules are essentially those of a semantic grammar (Burton 1976). The TAUM-METEO system (Chevalier et al. 1978) for translating Canadian weather bulletins is based on a grammar arrived at through a distributional analysis of a large corpus of these texts. In less stereotyped sublanguages such as medical records, there may be both elliptical sentence forms and their full-sentence paraphrases in the sublanguage. Thus the NYU system for extracting formatted data from medical records (Sager 1978, 1981) must include in its parser special rules for elliptical forms as well as more general syntactic rules for the full forms.

Most sublanguages of English observe the syntactic patterns of standard English but may differ markedly in the frequency of usage of various constructions. For example, many of the question forms, stylistic inversions, and exclamatives of conversational English are totally absent from technical literature (Lehrberger 1981). Grammars for processing technical language may therefore delete the corresponding production rules for analysis in technical domains. On the other hand, some sublanguages may use syntactic construc-

tions unknown in the general language, in which case the appropriate productions must be included in the sublanguage grammar.

Even when certain grammatical constructions cannot be ruled out of the grammar, they may be of such high or low frequency in the sublanguage that this fact can be used to reorganize the order in which rules are tried or to change the preference weighting assigned to competing syntactic analyses.

5.3.3 Semantic Constraints

The restricted domain of reference of a sublanguage is mirrored in the way words are used with respect to one another. A distributional analysis of word co-occurrences in a large corpus of texts (Harris 1963; Hirschman, Grishman, and Sager 1975) allows a computational linguist to group words into equivalence classes and to describe the occurring sentences in terms of these classes. Computational systems which use the semantic grammar approach (Burton 1976) state the syntax directly in terms of such distributional classes, which are relevant for the semantic or functional distinctions to which the system is sensitive. Collapsing syntax and semantics in this way is useful for small sublanguages (Hendrix et al. 1978; Epstein and Walker 1978), but there is the disadvantage that the grammar has no generality and a new one has to be written for each new sublanguage. Though one argument for semantic grammars has been that they are computationally more efficient, recent experiments in which a semantic grammar was compared with a linguistically motivated grammar for the same database demonstrated that the latter could be just as efficient (cf. Sagalowicz 1980).

In more complex sublanguages it is usually necessary to maintain traditional syntactic categories, and hence to couch parsing rules in terms of these categories. In this case, semantic constraints in the form of selectional restrictions can be applied either during or directly after parsing to eliminate those syntactic analyses that give meanings impossible in the sublanguage (Sager and Grishman 1975, Sager 1981, Robinson 1980).

Most sublanguage texts also have larger information structures beyond the word-class co-occurrences of single sentences. An analysis of the information formats of medical records (Hirschman and Sager 1981) has been carried out for the purpose of information retrieval. Frame-like structures may also be employed to recognize and extract larger information components (e.g., Bobrow et al. 1977, Schank et al. 1980).

A number of techniques are being developed for the specification and representation of semantic structures that can extend beyond the sentence unit. One

entails the assignment of propositional structures to text passages (Walker and Hobbs 1981). Domain and protocol analysis (Davis 1977, Newell and Simon 1972, Malhotra 1975) provide techniques for hypothesizing facts and inference rules appropriate for semantic analysis and reasoning procedures. Knowledge acquisition procedures (Davis 1977, Haas and Hendrix 1980, Rychener 1980), now under investigation, could significantly aid in the building of semantic and inference components.

5.3.4 Discourse Considerations

Recent research has shown that the way in which sentences are strung together to form coherent text can vary considerably from one sublanguage to another. In addition to differences in discourse-level semantic structures (see 5.3.3), separate sublanguages may make different use of a language's linguistic means of textual cohesion. In view of the considerable attention given to anaphora in the literature of computational linguistics, it is worthwhile to note that certain technical sublanguages contain no occurrences of anaphoric pronouns, while others make use of special anaphoric devices (Kittredge 1981). Even when a technical sublanguage uses pronominal anaphora, it often appears that the sublanguage effectively restricts it to cases where the antecedent noun phrase occurs in the preceding sentence or even in an earlier clause in the same sentence. Needless to say, the strategy employed for establishing co-reference in a sublanguage must therefore take into account the behavior of each anaphoric device in that same sublanguage. In many cases, a far simpler algorithm can be used than would be necessary for unrestricted language. In any given language, the semantic coherence and grammatical cohesion of a text can be signalled by a variety of linking devices. From a language's inventory of devices, each sublanguage seems to make a rather distinctive and limited selection. Stock market reports avoid repetition of the same verb in successive sentences, using synonyms instead, whereas technical manuals apparently avoid synonymy at the expense of lexical repetition (Kittredge 1981). The use of tense or tense variation may also fit a distinctive pattern. All such tendencies, whether probabilistic or absolute, may be exploited during the design of optimized sublanguage processing systems.

5.4 Factors Defining Suitable Candidate Applications

The sublanguage approach to language processing may not be appropriate to all varieties of restricted language or all types of application. It may only be profitable where there exists an established group of users who help to identify and define the knowledge domain. In addition, the domain should be relatively well-defined and internally consistent. The most tract-

able sublanguages from the computational point of view are those that present a simple discourse structure. Finally, each application should be one in which the computer is an appropriate medium of communication or processing (e.g., spoken sublanguages or ones for which permanent records would not or should not be kept may not be appropriate).

In practical applications where economic considerations are decisive, one must also take into account the time and cost of studying the linguistic properties in a sufficiently large and representative sample of the sublanguage and of creating and programming the sublanguage-specific dictionary and grammar rules. There is reason to believe that sublanguages that are semantically and pragmatically near-neighbors are similar in their grammatical properties, so that a better understanding of language form and function will make the description of new sublanguages easier and more predictable.

5.5 Maturing Areas of Research Relevant to the Sublanguage Approach

A successful general approach to sublanguage processing in a wide variety of domains will depend on advances in a number of research areas, some of which are maturing rapidly. Empirical work on knowledge structures (Bobrow et al. 1977, Mark 1980, Robinson et al. 1980) and on mechanisms of focus (Grosz 1977, 1981) is relevant to a proper treatment of sublanguage specific features of discourse and semantic structure.

Techniques of using precise selectional restrictions for sublanguages have been implemented (Burton 1976) as have those for extracting formatted information from fairly stereotyped sublanguages (Sager 1978). A new technique for developing transportable systems for natural-language interfaces to databases (Hendrix and Lewis 1981) elicits from the user a language for querying the contents at the same time that information about the domain is being entered. This approach is being extended to provide a more sophisticated system that is not limited to formatted databases but entails translation into a set of well-formed formulas in a many-sorted first-order logic (Haas and Hendrix 1980). Recent work on treating departures from grammaticality (Sondheimer and Weischedel 1980, Hayes and Mouradian 1980, McKeown 1980, Kwasny and Sondheimer 1981, Miller et al. 1981) can be used in handling specialized language that deviates syntactically from the standard language. Devices for designing more "friendly" systems, such as the work on graceful interaction (Kaplan 1978, Hayes and Reddy 1979, Weischedel and Sondheimer 1981) are relevant to the question of relating sublanguage-specific phenomena to those of the whole language.

5.6 Promising New Research Areas

A number of new or even underdeveloped research areas will certainly prove important for work on sublanguage. We expect that further research on syntactic variation will yield a more unified framework for the description of sublanguage word and phrase structure. Work in pragmatics, such as the recent computational modeling of speech acts, will intersect with investigations into sublanguages where social or legal dimensions are important. As we accumulate experience in semantic processing over a number of specialty areas, we will be able to identify more and more sharply the important parameters for assessing the computational tractability of any given sublanguage. This experience will also nourish a distinct area which has both theoretical and practical aspects: the problem of relating sublanguages (and their grammars) to the standard language (and its grammar). The preliminary efforts at building up a taxonomy and typology of sublanguages are aimed in this direction.

There is already an identifiable movement towards codifying and teaching language for specific purposes. For some applications it is possible to take naturally occurring sublanguages and slightly regularize them so that strong tendencies are promoted to norms for communicating in the subfield. Attempts in this direction have occurred in the stylistic guidelines now used for writing weather reports and aircraft maintenance manuals. A serious scientific approach to this "engineering design" of new sublanguages must await a more exact theoretical and practical understanding of how language function relates to language form.

5.7 Recommendations

At present, only a small number of sublanguages have been studied in detail. Thus one urgent need is to broaden the basis of our understanding of these linguistic subsystems. The members of the panel feel that this can best be achieved by selecting a few promising application areas in which to concentrate substantive research resources. Such concentration is necessary for several reasons. First, most naturally occurring sublanguages present real challenges for linguistic description. Many months or years of effort must usually be invested in describing a corpus of texts and in finding the natural extensions of that corpus in collaboration with speakers of the sublanguage. Second, the linguistic peculiarities of the sublanguage often present new problems for computational treatment, particularly if the solutions are to be generalizable to other, related sublanguages. Third, many further months of on-site testing are usually necessary to properly absorb and evaluate the feedback from users of prototype systems, and to evolve more adequate versions. The evolution of any significant new system therefore implies a substantial collaborative effort over a period ranging from several months to several years.

In parallel with a program of applied research along the lines suggested above, we recommend that certain kinds of basic research be supported which can both feed and be nourished by the applied research. Basic research in the areas identified under Sections 5.5 and 5.6 above should be encouraged in such a way that researchers, however theoretically oriented, are brought periodically into contact with the practical aspects of the proposed real-world applications. Such an interplay between the practitioners of basic and applied research has proved to be an essential ingredient of past advances in sublanguage processing.

6. Acknowledgements

This workshop is the first in a series organized by the Navy Center for Applied Research in Artificial Intelligence. The concept for this workshop emerged from numerous discussions with Marvin Denicoff and Joel Trimble of ONR, Paul Chapin and Henry Hamburger of NSF, Robert Engelmores and Robert Kahn of DARPA, Stanley Wilson and John Davis of NRL, and William Price of AFOSR.

The workshop itself was made possible only through the superb cooperation of the ACL. Norm Sondheimer, former ACL president, and Don Walker, ACL Secretary-Treasurer, used their organizational talents to incorporate the workshop into the 1981 ACL Conference. Jerry Kaplan, local chairman for the ACL meeting, graciously accepted the added responsibility of providing local arrangements for the workshop.

We gratefully acknowledge the very competent secretarial assistance by Janet L. Stroup of NRL and the careful compilation of the workshop proceedings by Veronica Bates of NRL. Financial support for the workshop was provided by the Office of Naval Research.

References

- Artificial Intelligence Corporation 1981 *Intellect Query System User's Guide*. Release 101. Artificial Intelligence Corp., Waltham, MA.
- Biermann, A. and Ballard, B. 1980 Toward Natural Language Computation, *AJCL* 6 No. 2, 71-86.
- Bobrow, R. 1978 *The RUS System*. BBN report 3878. Bolt, Beranek, and Newman, Inc., Cambridge, MA.
- Bobrow, D., Kaplan, R., Kay, M., Norman, D., Thompson, H., and Winograd, T. 1977 GUS, A Frame-Driven Dialogue System. *Artificial Intelligence* 8 155-173.
- Brown, J., Burton, R., and Bell, S. 1974 *SOPHIE: A Sophisticated Instructional Environment for Teaching Electronic Troubleshooting*. BBN Report 2790 (March).
- Burton, R. 1976 *Semantic Grammar: An Engineering Technique for Constructing Natural Language Understanding Systems*. BBN Report 3453, Bolt, Beranek, and Newman, Inc. Cambridge, MA (December).
- Carbonell, J. and Hayes, P. 1981 *Dynamic Strategy Section in Flexible Parsing*. Nineteenth Annual Meeting of the Association for Computational Linguistics. Stanford, CA (June).
- Chevalier, L., Dansereau, J., and Poulin, G. 1978 *TAUM-METEO: Description du Système*. Université de Montreal, Canada.
- Codd, E. 1974 Seven Steps to Rendezvous with the Casual User. In Klimbie, J. and Koffeman, K., Eds., *Data Base Management*. North-Holland, Amsterdam: 179-200.
- Codd, E. 1978 How About Recently? (English Dialogue with Relational Databases Using RENDEZVOUS Version 1). In Shneiderman, B., Ed., *Databases: Improving Usability and Responsiveness*. Academic Press, New York: 3-28.
- Davis, R. 1977 Interactive Transfer of Expertise: Acquisition of New Inference Rules. In *Proceedings of the Fifth International Conference on Artificial Intelligence*. Cambridge, MA: 321-328.
- Eppstein, M. and Walker, D. 1978 Natural Language Access to a Melanoma Data Base. In *Proceedings of the Second Annual Symposium on Computer Application in Medical Care*. IEEE, New York: 320-325.
- Grishman, R. and Hirschman, L. 1978 Question Answering from Natural Language Data Bases. *Artificial Intelligence* 25-43.
- Grosz, B. 1981 Focusing and Description in Natural Language Dialogues. In Joshi, A., Sag, I., and Webber, B., Eds., *Elements of Discourse Understanding: Proceedings of a Workshop on Computational Aspects of Linguistic Structure and Discourse Setting*. Cambridge University Press, Cambridge: 84-105.
- Grosz, B. 1977 The Representation and Use of Focus in a System for Understanding Dialogs. *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*. Cambridge, MA (August 22-25): 67-76.
- Haas, N. and Hendrix, G. 1980 An Approach to Acquiring and Applying Knowledge. In *Proceedings of the First Annual National Conference on Artificial Intelligence*. Stanford University: 235-239.
- Hendrix, G. and Lewis, W. 1981 Transportable Natural-Language Interfaces to Databases. In *Proceedings of the Nineteenth Annual Meeting of the Association for Computational Linguistics*. Stanford, CA (June).
- Hendrix, G., Sacerdoti, E., Sagalowicz, D. and Slocum, J. 1978 Developing a Natural Language Interface to Complex Data. *ACM Transactions on Database Systems* 3 No. 2 (June) 105-147.
- Harris, Z. 1963 *Discourse Analysis Reprints*. The Hague, Mouton.
- Hayes, P. and Mouradian, G. 1980 Flexible Parsing. In *Proceedings of the 18th Annual Meeting of the Association for Computational Linguistics*. University of Pennsylvania: 97-103.
- Hayes, P. and Reddy, R. 1979 Graceful Interaction in Man-Machine Communication. *Sixth International Joint Conference on Artificial Intelligence*. Stanford University, 372-374.
- Hirschman, L., Grishman, R., and Sager, N. 1975 Grammatically-Based Automatic Word Class Formation. *Information Processing and Management* 11.
- Hirschman, L. and Sager, N. 1981 Automatic Formatting of a Medical Sublanguage. In Kittredge and Lehrberger.
- Joshi, A., Mays, E., Lanka, S., and Webber, B. 1981 Natural Language Interaction with Dynamic Knowledge Bases: Monitors as Responses. In *Proceedings of the IJCAI 1981*. Vancouver, Vancouver (August).
- Kameny, I. et al. 1978 An End User Friendly Interface for Databases. *Proceedings VLDB*. Berlin.
- Kaplan, S.J. 1979 *Cooperative Responses from a Portable Language Data Base Query System*. Ph.D. Dissertation, University of Pennsylvania. (Available as Stanford Heuristic Programming Project Report HPP-79-19, Computer Science Department, Stanford University, Stanford, CA, 94305 (July).)
- Kaplan, S.J. 1978 Indirect Responses to Loaded Questions. *Theoretical Issues in Natural Language Processing-2*. University of Illinois at Urbana-Champaign (July).
- Kaplan, S.J. and Davidson, J. 1981 Interpreting Natural Language Database Updates. In *Proceedings of the Nineteenth Annual Meeting of the Association for Computational Linguistics*. Stanford, CA, June.
- Katter, R. and Montgomery, C. 1972 On-Line Bugging: Hope for Terminal Cases of Semantic Deviance. Invited Paper at the Gordon Research Conference on the Frontiers of Science. New London, NH, July.
- Kay, M. 1979 Functional Grammar. In *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society*.
- Kittredge, R. 1981 Variation and Homogeneity of Sublanguages. In Kittredge and Lehrberger.
- Kittredge, R. and Lehrberger, J. Eds. 1981 *Sublanguage: Studies of Language in Restricted Semantic Domains*. deGruyter, Berlin.
- Kwasny, S. and Sondheimer, N. 1981 Ungrammaticality and Extragrammaticality in Natural Language Understanding Systems. In *Proceedings of the Seventeenth Annual Meeting of the ACL*. La Jolla, CA (August).

- Landsbergen, S. and Scha, R. 1978 Formal Languages for Semantic Representation. In Petofi, J., Ed., *Aspects of Automated Text Processing*. Buske, Hamburg.
- Lehrberger, J. 1981 Automatic Translation and the Concept of Sublanguage. In Kittredge and Lehrberger.
- Mckeown, K. 1980 Generating Relevant Explanations: Natural Language Responses to Questions About Database Structure. In *Proceedings of First Meeting of AAAI*. Stanford, CA (August).
- Malhotra, A. 1975 Design Criteria for a Knowledge-Based English Language System for Management: An Experimental Analysis. MAC TR 146. Cambridge, MA: Project MAC. Massachusetts Institute of Technology, February.
- Mark, W. 1980 Rule-Based Inference in Large Knowledge Bases. In *Proceedings of the First Annual National Conference on Artificial Intelligence*. Stanford, CA.
- Mays, E. 1980 Failures in Natural Language Systems: Applications to Database Query Systems. In *Proceedings of the First Meeting of AAAI*. Stanford, CA, August.
- Miller, L., Heidorn, G., and Jensen, K. 1981 Text-critiquing with the EPISTLE System: An Author's Aid to Better Syntax. In *AFIPS Conference Proceedings*. AFIPS Press, Montvale, NJ: 649-655.
- Moore, R.C. 1981 Problems in Logical Form. In *Proceedings of the Nineteenth Annual Meeting of the ACL* (June).
- Newell, A. and Simon, H. 1972 *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.
- Novak, G.S., Jr. 1981 Physics Problem Solving: ISAAC-II. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*. IJCAI-81, Vol. 2. University of British Columbia, Vancouver, B.C. (August).
- Petrick, S. 1975 Design of the Underlying Structure for a Data Base Retrieval System. In Grishman, R., Ed., *Directions in Artificial Intelligence: Natural Language Processing*. Courant Computer Science Report 7, Courant Institute of Mathematical Sciences, York University, New York, New NY, 60-93.
- Robinson, A., Appelt, D., Grosz, G., Hendrix, G., and Robinson, J. 1980 Interpreting Natural Language Utterances in Dialog about Tasks. *Communications of the ACM* in press. SRI Technical Note 210. Artificial Intelligence Center, SRI International, Menlo Park, CA.
- Robinson, J. 1982 DIAGRAM: A Grammar for Dialogues. *Communications of the ACM*.
- Rychener, M. 1980 Approaches to Knowledge Acquisition: The Instructable Production System Project. In *Proceedings of the First Annual Conference on Artificial Intelligence*. Stanford, CA: 228-230.
- Sagalowicz, D., Ed. 1980 Mechanical Intelligence: Research and Applications. Final Technical Report. Artificial Intelligence Center, SRI International. Menlo Park, CA.
- Sager, N. 1981 *Natural Language Information Processing: A Computer Grammar of English and Its Applications*. Addison-Wesley, Reading, MA.
- Sager, N. 1978 Natural Language Information Formatting: The Automatic Conversion of Texts to a Structure Data Base. In Yovits, M. and Rubinoff, M., Eds., *Advances in Computers*. Academic Press, New York: 89-162.
- Sager, N. and Grishman, R. 1975 The Restriction Language for Computer Grammars. *Communications of the ACM* 18 390-400.
- Schank, R., Lebowitz, M. and Birnbaum, L. 1980 An Integrated Understander. *American Journal of Computational Linguistics* 6 13-30.
- Silva, G.M.T., Dwiggin, D.L., Busby, S.G., and Kuhns, J.L. 1979 *A Knowledge-Based Automated Message Understanding Methodology for an Advanced Indications System*. OSI Report R79-006, Operating Systems, Inc. (February).
- Simmons, R. F. and Chester, D. Relating Sentences and Semantic Networks with Clausal Logic. *Communications of the ACM*, to appear.
- Sondheimer, N. and Weischedel, R. 1980 A Rule-Based Approach to Ill-Formed Input. In *Proceedings of the Eighth International Conference on Computational Linguistics*: 46-53.
- Thompson, F. and Thompson, B. 1975 Practical Natural Language Processing: The REL System as Prototype. In Rubinoff, M. and Yovits, M.C., Eds., *Advances in Computers, Volume 13*. Academic Press, New York.
- Walker, D. and Hobbs, J. 1981 Natural Language Access to Medical Text. In *Proceedings of the Fifth Annual Symposium on Computer Applications in Medical Care*. IEEE, New York.
- Waltz, D.L. 1978 An English Language Question Answering System for a Large Relational Data Base. *Communications of the ACM* 21 526-539.
- Weischedel, R. and Black, J. 1979 *Responding to Potentially Unparsable Sentences*. Tech Rep. 79/3. Department of Computer and Information Sciences, University of Delaware, Newark, DE.
- Weischedel, R. and Sondheimer, N. 1981 *A Framework for Processing Ill-Formed Input*. Technical Report. Department of Computer and Information Sciences, University of Delaware, Newark, DE.
- Wilensky, R. 1978 *Understanding Goal-Based Stories*. Yale University Research Report No. 140.