TEUCHISTIC NATURAL LANGUAGE PROCESSES

Robert K. Lindsay
The University of Michigan, Ann Arbor, Michigan 48109

Alexis Manaster-Ramer
Wayne State University, Detroit, Michigan 48202
and
IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

Keywords: natural language, teuchistic

ABSTRACT

AI approaches to natural language specify computational processes, yet they are based on the structural concepts of language and grammar which posit necessary conditions at least for the "correct" interpretations of utterances and often also for the syntactic and/or logical representations of utterances. We argue that any such view will fail to account for a variety of important features of language behavior, which we describe. Systems based on the language/grammar model are usually accompanied by heuristic or algorithmic search/selection methods of computation. We contrast these methods with constructive (*teuchistic*) computation models based on redundant, inconsistent sets of constraints, and show that this view offers natural accounts of the phenomena described.

I. INTRODUCTION

Almost all computational work on natural languages, in and out of AI, has adopted the conception of language, derived from traditional grammar and structural linguistics, according to which there exists a body of knowledge which defines the primitives of which the language is composed, the principles by which these primitives may be combined, and the meanings associated with each primitive and with each principle of composition. This conceptual homogeneity has been obscured by the controversies surrounding virtually every other question of NL processing, such as the debate over models which postulate syntactic and/or logical levels of analysis in addition to the meaning (conceptual) representation (e.g., LUNAR (Woods *et al.* 1972), SHRDLU (Winograd 1972), PARSIFAL (Marcus 1978)) as opposed to models which relate utterances to meanings in an integrated fashion (e.g., Wilks's (1975) parser, ELI (Riesbeck and Schank 1976), and the Word Expert Parser (Small 1980)),[1] which has sometimes been viewed as involving a contrast between grammar-based and knowledge-based methodologies[2]

We are thus abstracting from the controversies about exactly how this body of knowledge is structured and whether it should be used—in parsing, say—together with world knowledge. Rather, we are concerned with the fact that almost all models purport to simulate the idealized user of a given language, conceived of as an expert on his native language (e.g., Cullingford 1986:34), and

---

[1]The issues separating these approaches have at any rate been settled to some extent at least by the compromise position of Lytinen (1986), who proposes that syntactic and semantic information form distinct knowledge bases but that these are combined dynamically during parsing.

[2]A related issue concerns the distinction, if any, between a declarative representation of knowledge of a language and a processing model. However, since in models that contain little or nothing beyond a processor, the latter still implicitly represents a description of the language, this distinction will also be ignored in this paper. After all, a context free grammar could be presented by means of a process model such as a pushdown automaton or an Earley parsing algorithm.

consequently assume that there is such a thing as a definite knowledge of *what* a native speaker does in the way of assigning some structural representations.[3]

We thus wish to contrast such **models of language** with **models of (potential) linguistic behavior**, which assume that the meanings (and any intermediate representations) are arrived at by means of ***constructive*** processes which take into account a variety of factors, only some of which bear a close resemblance to the linguistic knowledge as normally conceived, whereas others involve factors such as the past processing that has been done by the system, interaction with other users of the language, and even physical characteristics of the language user.

We would discard the concepts of ***language*** and ***grammar***, replace the customary notion of linguistic competence with a concept of ***potential behavior***, and view production and comprehension as nonmonotonic problem-solving processes that satisfice rather than search out an optimal choice from a pre-established theoretical set of possibilities. We further suggest that the multi-component constraint satisfaction model serve as the basis of such problem solving. In this paper we outline the approach without specifying a particular theory, and we refer to Lindsay (1971) as an example of this approach in application. We will discuss some of the principal kinds of facts of human linguistic behavior which justify and perhaps even necessitate this change of perspective. We will show that this approach is closely related to satisficing models of economic behavior and to nonmonotonic models of reasoning, while the conventional language-based models are essentially optimizing and monotonic.

II. TEUCHISTIC MODELS OF LINGUISTIC BEHAVIOR

Teuchistic Computation

Typically, AI problem solving methods involve a search through a problem space or game tree. Although there may be an exact algorithm involved in simple cases, usually the process is heuristic.[4] Our view is that NL use involves processes which are neither algorithmic nor heuristic but, to coin a new term, **teuchistic**. Teuchistic processes construct a solution rather than search for one through a set of pre-established possibilities. Trivially, many numeric computations are of this sort. For example, squaring a number (or evaluating any other closed algebraic form) is a process of constructing the answer rather than searching through the set of possibilities (the real numbers, in this case).[5]

Less trivially, at times it is not possible to define a solution space extensionally as when the space is infinite, but one can do so intensionally (for example, the real numbers). Furthermore, even when such a definition is possible, it may not be enlightening if the computation itself is not cast as a search within that solution space. Consider the case of design, say of houses or electronic circuits. Here computation is best viewed not as search but as construction, for any definition of the set of possible houses or circuits will not aid the design process as it is usually carried out by human designers. Rather the process is one of constructing a solution using available tools and meeting a set of heterogeneous constraints. Indeed if the constraints are not themselves necessary conditions, but merely design desiderata ("suggestions"), and if further the constraints may be vague and inconsistent (as design

---

[3]We will also ignore such important and controversial questions as the following: Should a computational model simulate *how* human beings represent linguistic knowledge? What level of understanding should be aimed at? Should pragmatics (knowledge of the world) be used to guide NL understanding? Will models designed for small domains of discourse scale up to whole languages?

[4]Some work in computational linguistics, in fact, views human linguistic processing as involving a highly inexact heuristic, which can handle only a restricted subset of the language allowed by the theoretical model of linguistic representation (e.g., Peters, to appears).

[5]However, some numeric computations involve search, usually systematic; for example computing a square root by Newton's method or any other algorithm that works by successive approximation may be seen as as search process.

principles generally are), the concept of a solution space is not even coherent. In such situations, as we will argue is the case for natural language use, computations must be teuchistic; further, any attempt to circumscribe the set of permissible "solutions" will be futile and a diversion from the real task.

As we will argue later, the conventional approaches to NLs appear to be optimizing in that they search for the "correct solution" to questions of syntactic or semantic representation of a particular utterance and monotonic in that the system of rules defining a correct solution is self-consistent and allows the representations to be deduced following the rules of classical logic. It should be emphasized that this is trivially true even for those models which have no syntactic or logical representations but which construct a meaning representation directly from the utterance and the knowledge of the world. The teuchistic approach, on the other hand, shares important features of satisficing and nonmonotonic models which have been developed within AI in domains other than NL.

From our point of view, the crucial distinction is between models that assume that there is a thing called language (specifically a particular language at a particular time, e.g., late 20th century American English) which can be usefully modeled computationally, usefully in the sense that such a model would behave sufficiently like a human native speaker of this language to be able to replace human beings at some significant tasks but also in the sense of casting new light on the workings of human intelligence.

Our model, if it is to be functional at all, almost necessarily has to be redundant and multipartite. This view assumes the existence of several different *sources of constraints*, each of which can be viewed as a generative component. It is not necessary that these sources be non-overlapping, consistent, nor even clearly categorized. It is only necessary that they be *redundant* in the sense that structural-interpretive information available from one set of sources may be used to infer reliably (though not necessarily with certainty) structural-interpretive information of other types. For purposes of illustration, we may assume sources of syntactic, lexical, phonetic, semantic, and pragmatic constraints (but this is an oversimplification); in addition there may be sources of logical and empirical constraints of various kinds.

It is the availability of these multiple sources of redundant constraints that drives, for example, the comprehension process. Comprehension is modelled as a constraint driven satisficing process: information is selectively attended to and used to constrain the generation of an interpretation or utterance. If all structural information is consistent, the interpretation is rapid and unproblematic. If some structural information is not consistent with the interpretation under construction, a strategy of backtracking is undertaken to revise a previous decision. If backtracking fails (outright inconsistency is encountered), some constraint must be relaxed, again according to a specific strategy.

A specific earlier proposal along these lines involved learning mechanisms based on a multiple components model which has been described and programmed by Lindsay (1971). The learning mechanism is called the *jigsaw heuristic*,[4] after the jigsaw puzzle. A jigsaw puzzle provides two sources of constraints for its solution: information from the contours (syntax?) that determines what pieces may be adjacent, and information from the design (semantics?) about the color and form of the scene depicted. These sources are redundant in the sense that the puzzle could be done if all the pieces were congruent squares *or* if all the pieces were upside down. Now suppose some of this information were incomplete – tabs were deformed, or coffee spilled on the design. Solving the puzzle would then lead to clear hypotheses about the nature of the missing information. In the NL situation, missing knowledge of word meaning or part-of-speech could be induced by a learner if sufficient information remained to permit the construction of a satisfactory interpretation.

Modeling Potential Behavior

In place of language, viewed as an object about which a definite body of information is available, we choose to focus on human linguistic behavior. The distinction is, as noted, not reducible to that

---

[4]Though we now see a crucial difference between our approach and the usual heuristic models.

between declarative grammar formalisms and processing formalisms; nor is it the same as the competence/performance distinction as formulated by Chomsky and others. Perhaps an analogy will make this clear. Human athletic ability is limited by the laws of physics; it is further limited by the specific configurations of bone and muscle and by physiological constraints; it is still further limited at any particular time and place by the exigencies of the context, the conditioning of the athlete, the weather, equipment, and so forth. To say that human athletic competence is limited only by the laws of physics, and that all other limitations (such as inability to jump over the moon) are "performance" limitations, is not a theory that in any way tells us about human beings.

On the other hand, a theory of athletic ability that specified general limitations due to human anatomy and physiology would be highly informative, even though it did not enable the prediction of a specific athlete's performance on a given day. But such a theory must tell us about the range of *potential behavior*, specifically, it must delimit what *could actually be done under the proper conditions*. Thus it would need to make reference to such factors as muscle strength, blood supply, and so forth, and how these control processes.

The theory we seek is a theory of *potential behavior* in this sense.[7] That is, we are not interested in what a specific speaker may or may not have done, nor in attempting to predict what he will or will not do in the future. Rather, we are interested in predicting what he might do under the proper conditions. It seems to us that there are a large number of claims along these lines that are essentially implicit in previous scholarship, although they have rarely been formulated in this way. There is in existence a philosophical literature (e.g. Itkonen 1978), which points out that linguistic theories are not theories of the same order as those found in other sciences precisely because they lack the ability to make these kinds of if-then claims. It seems to us that the misunderstandings about competence and performance lie at the root of the problem. One cannot make verifiable if-then claims about competence, which is by definition inaccessible, and no serious claims could be attempted for actual performance. It is only by focusing on potential, idealized *performance* (behavior) that we are able to make these kinds of claims.

III. LIMITATIONS OF LANGUAGE-BASED MODELS

What is a Language-Based Model?

Despite the diversity of fields interested in language and the diversity of approaches to its study, there is a remarkable unanimity about the existence and the boundaries of this phenomenon. The reason for this lies in the general acceptance of certain notions sometimes taken to be pretheoretical constants, but which are in fact largely the product of centuries of scholarly and other activity on language in the Western culture. Typical of this general consensus are the following: there are many languages, each of which is a separate entity; languages consist of sounds which make up words which make up sentences; hence, for each language, it is possible to set up an analysis that tells us just what it consists of; such an analysis will take the form of several chapters of a grammar plus a lexicon; this analysis represents a kind of knowledge, knowledge of one's language, that is distinct from other knowledge, such as knowledge about the physical world.[8]

The heart of the paradigm is this assumption that language can be studied as an entity in its own right. It is generally presupposed that it *is* possible to draw a boundary between knowledge of language and other knowledge. AI work which uses world knowledge nevertheless assumes that there *is* a body of knowledge about, say, modern American English which is self-contained. Thus, most AI models have been based on some conception of grammar, either explicitly by use of grammar rules plus

---

[7]We believe that a case can be made that such an approach was envisaged by the students of language known as neogrammarians at the turn of the century (e.g., Paul 1880).

[8]Even if we assume, as many AI researchers do, that knowledge of the real world guides—or assists in—the processing of natural language.

parser or implicitly by virtue of their reliance on representations of linguistic knowledge, whether in declarative or in process terms, which express the same information as a grammar.

The concept of grammar naturally suggests a model of language understanding that employs a grammar to assign structural descriptions to sentences, and then computes meanings from these structural descriptions. Research in AI, especially in the early days of NL understanding, has commonly been based on such a model. A major step in such a research program is of course to write a grammar to characterize the set of permissible sentences, and this task is formidable. However, it is not the worst of the problem. As soon as a grammar of any size is used in an analysis system, it is discovered that the "parser" produces a surprising number of structural descriptions for even very short and seemingly simple sentences. Initially, this problem was thought to be a minor inconvenience that caused very long computational times, but which could be gotten around in other ways. In fact the problem is deeper, since very large numbers of structural descriptions are the rule, not the exception. So with the absolute in-or-out criterion of a grammar, one is caught on the horns of a dilemma: either the generative capacity is severely limited, or every sentence becomes multiply ambiguous to an extreme, and no explanation is offered for the fact that humans fail to even *notice* most of the possible readings, though they can do so with effort, or at least understand explanations of them.[9]

On the other hand, more recent AI work has chosen to analyze just fragments of a language and has focused on the use of world knowledge to guide the process of arriving at an interpretation of a sentence. This approach necessarily fails in face of the same facts, though for the opposite reason. In these approaches, there must be an arbitrary dividing line drawn between those interpretations which are "normal" enough to be allowed by the world knowledge and those which will never be allowed. Yet, the observed fact is that people do arrive at readings which no AI program would allow. In fact, the reason that AI researchers have been so conscious of the need to use world knowledge in this way is precisely because, qua human users of English, we are able to arrive at all sorts of unlikely readings, which we then need to rule out.[10]

In sum, the uses to which humans put language are strongly dominated by many considerations, only some of which could be expressed as conventional linguistic knowledge (whether in terms of a grammar or of a heuristic understanding program). On the one hand, a language is more than a set of pairings of utterances and "conceptual representations."[11] On the other hand, the linguistic behavior of human beings is not restricted to some tiny set of "normal" sentences with "normal" interpretations. What interpretations will accrue to any utterance (and, likewise, how a given meaning will come to be expressed) depends not only on the supposed knowledge of the language but on the previous linguistic experiences of the language user (human or simulated), on external circumstances, and very likely on physical and other relatively low-level characteristics of the user.[12]

---

[9]While we are well aware of the proposals that human beings employ processing mechanisms which can handle only a small part of what the grammar generates, very little has actually been done to show that such a model can account for the facts. Indeed, as we point out in the next footnote, there is a strong reason for doubting that this could be the case.

[10]This argument applies with equal force to grammar-plus-processor models which claim that people can process only a subset of the language generated by the grammar, since the dividing line between what is processable ('acceptable' in the standard terminology) and what is not is equally arbitrary and empirically unmotivated.

[11]A fortiori, we reject any view of a language as a set of uninterpreted sentences.

[12]In this context, the most obvious phenomena are the properties of language which are due to the neurophysiology of articulation, the physics of air-borne sounds, and the like, as noted by Paul (1880). In the area of syntactic processing, the most striking example of a hypothesis along these lines would be Yngve's (1960) proposal that human beings use pushdown stacks of extremely limited depth.

Arguments for the Teuchistic Approach

We now turn to some general patterns of linguistic processing by human beings which seem to support a teuchistic model.

- People produce and parse rapidly a vanishingly small proportion of the theoretically possibly utterances of a given language, and moreover, intend and understand only a vanishingly small proportion of the theoretically possible interpretations of those utterances.
- Yet, given sufficient time and effort, people seem to be able to increase these proportions significantly, which is precisely why we are aware of the theoretical possibilities in the first place (in a real sense, they are not merely theoretical).[13]
- It is obvious that any written or spoken sequence can be used as a proper name of a person, place, piece of music, art, or writing, musical group, political party, etc. Such strings can moreover have the form of sentences or parts of sentences which are otherwise "impossible". Thus, ***Do you know why you don't not say no sentences like this?*** would be ruled out as ungrammatical by some NL systems and forced to mean something like ***Do you know why you don't say any sentences like this one*** by others, yet it can, and for some people is likely to, mean something like ***Do you know [McCawley's (1973) article] "Why you don't not say no sentences like this one"?***. Thus, Charniak (1983: 118) is wrong in assuming that ***The boys is dying*** must be either left uninterpreted or else (as in his system, PARAGRAM) interpreted as ***The boys are dying***. If ***The boys*** were the name of an avantgarde musical group, and perhaps it is, then this sentence would be interpretable without any such coercion. It follows that in a sense the task of generating all and only grammatical sentences is trivial: generate everything. What will be nontrivial will be the task of constructing appropriate interpretations for certain sentences, namely, those which can most naturally be interpreted by taking certain parts of them as (unusual) proper names.[14]
- A related difficulty, noticed for example by Robinson (1984) in work on DIAGRAM and by many others, is that common nouns (and other categories of words) can be—and are—created often enough to make reliance on any fixed dictionary an impossibility for a NL program of any scope.
- A similar but even deeper problem arises because it is possible to introduce arbitrary codes into a linguistic context. We might say, for example, that from now on ***They are flying planes*** is to be taken to mean that the phone is being tapped. The correct analysis of the code then depends on the interpretation of its definition at a point that is arbitrarily far removed from its use; this would require the grammar to be arbitrarily large to manage this feat. Alternatively, a grammar could trivially assign all possible interpretations to all possible strings. However, in the real-world codes are both finite and used under restricted circumstances, so the problem appears to be amenable to a teuchistic approach.

A more recent claim is that in favor of people using queues rather than pushdown stacks in linguistic processing (Manaster-Ramer, 1986).

[13]An extreme example involves expert informants, such as highly trained linguists, changing their minds over periods of years about the grammatical status or the interpretations of even short example sentences (e.g., Chomsky (1982) on parasitic gaps, Langendoen (1977; personal communication) on subject-verb agreement in *respectively* constructions). In such cases, NL users appear to operate in times often measured in years and not in milliseconds.

[14]Amsler (1987) has observed that a very large portion of naturally occurring text, such as news service reports, contains words not in dictionaries. Most of these are proper names that may be expressed in a wide variety of manners, so many that Amsler suggest the need for proper name formation rules rather than lexicons to record them; our view is that such methods will themselves be insufficient still. The same difficulty with proper names was encountered in a parser for Chinese character text developed by AMR's students as a class project.

• In real life, there are more competent readers than competent writers, more people adept at comprehension than articulate speakers. Language-based models in principle fail to allow for the possibility of such asymmetries—or else would consider it entirely outside their bounds. Models of behavior are called for.
• Equally striking differences separate different speakers, who despite different linguistic abilities, styles, dialects, and even (especially) knowledge of the topic are able to communicate to a degree, often a substantial degree. Again, language-based models offer no insight into this phenomenon, since they would have to simply recognize as many different languages as there are speakers. One argument against this is the fact that people recognize that certain individuals are better speakers of the *same* language than others (including themselves, oftentimes). All this suggests that we cannot seriously view the idealized (and idolized) native speaker as an expert whose knowledge is to be embodied in the NL programs.
• Communication can persist in the face of disruptions that cause deviations from *any* fully prescribed rules of well-formedness: typographical errors, acoustic distortions, neologisms, jargon, metaphor, halting and restarting, fractional sentences, and so forth. On the other hand, certain other types of interference, such as delaying the feedback of a speaker's own voice to his ears, is devastating. The latter fact is particularly revealing since it argues for a model that pays attention to the physical and other low-level properties of speakers.
• There are many cases where the language appears to change with use, i.e., the question of whether some utterance or utterance-interpretation pairing is licensed by a particular language cannot be answered in the abstract. Furthermore, the linguistic abilities even of a single speaker change constantly, even after puberty. Language-based theories could only deal with such facts by postulating ever newer grammars, which amounts to treating an evolving individual as a sequence of different individuals. Again, to the extent that such change is influenced by factors such as interaction with other speakers or one's physiological development, it falls outside the scope of language-based theories. The same is true, for similar reasons, of creative uses of language, such as metaphor, humor, introduction of words and terminology, and others.
• Perhaps the most striking phenomenon which argues for a teuchistic model is that of language change over generations. The existing models of language inevitably treat any language change, even a small one, as producing (or strictly speaking as being produced by) the rise of a new grammar, discontinuous from the old one. On our view (see also Paul 1880), language change is due to the fact that normal language processing involves feedback, so that the fact that particular forms were processed influences the underlying system. Since the processing is teuchistic and uses various subsystems, it follows that forms that are "incorrect" in terms of the overall system will actually occur and then in turn cause the system to be altered.[15] Moreover, since the overall system need not even be consistent, for many forms it may not even be possible to determine whether they are "correct" or not, yet they will occur, tending to restructure the system in the direction of making them "correct".[16]

All the phenomena discussed above seem to bear directly on the use of language in real-world settings, and increasingly problems due to them are becoming recognized in computational work, especially within the AI tradition. The problem of potentially unbounded ambiguity and difficulties due to the open-ended nature of names and even common words are examples which several researchers have had to deal with. As more and more ambitious system are attempted, more of these problems will be noticed to have debilitating practical effects. We believe that many of these difficulties can be solved by concerted and multi-pronged efforts of a teuchistic nature.

---

[15]The conventional view — even when supplemented with a feedback mechanism — leads to no such insights, since the it guarantees that every form produced will always be "correct" in the original grammar.

[16]For example, if a form occurs a sufficient number of times for whatever reason, it will tend to be reinforced. For example, people will often assimilate forms inconsistent with their original usage if these forms are frequently used by others in their presence (compare the use of infantile or pseudo-foreign forms by adults).

IV. PARALLELS TO OTHER WORK

There exist research paradigms in and out of AI that focus on analogous properties of nonlinguistic human behavior. Two examples are AI models of (nonmonotonic) reasoning and economic models of (satisficing) behavior. Both of these properties feature prominently in our ideas about natural language. The distinction between *optimizing* and *satisficing* processes, introduced by Simon (1947, 1955), illuminates the difference between the search-based models of NL understanding and teuchistic ones. Presumably most human problem solving involves satisficing—finding an adequate job, home, spouse—rather than optimizing—considering and comparing all jobs, houses, people—simply because information and computational resources are limited. On the view of language here proposed, the process of understanding an utterance never reaches completion in the sense of achieving *the* correct reading, nor does it ever consider all possible readings, nor is it necessary that such a set of possible readings exists. Rather, a sentence is interpreted in the context of a particular situation and purpose, so that changes in these factors could change the reading that is *constructed*. The teuchistic process continues until some adequate solution is achieved, or until one gives up the attempt for lack of time, knowledge, memory, or computational capacity. The end to be achieved may be to understand well enough to paraphrase, to answer a specific question, to give the appearance of comprehension, to determine the intelligence or political stance of the speaker, to memorize verbatim, and so on; it is not to find *the* meaning representation. An utterance is not a message to be decoded, it is a set of clues to a resolution of a problem.

The satisficing view that we have outlined leads to a potential connection between our proposals and the recent work in AI that has focused on the nonmonotonic character of human reasoning (see Turner 1984 for a survey). It would seem that most kinds of linguistic models lend themselves to a nonmonotonic manipulation. As far as the oldest attested more or less explicit grammar that has survived into our times (Pāṇini, no date, but definitely B. C.), we find that grammars contain mutually conflicting stipulations such that deductions based on two different rules (or, more generally, subgrammars) would often yield contradictory consequences. Normally, of course, deductions from a single rule of a grammar or a proper subset of the rules are not allowed. To see what a grammar generates, we have to consider the entire grammar, as befits an optimizing model. Hence, this nonmonotonic character of grammars has not had much significance. However, what we are proposing is that the normal situation in human linguistic behavior is satisficing, and this would naturally mean that, if there is such a thing as a grammar, only parts of it will be accessed at any given time in real-time processing. As a result, we would be able to explain how a speaker can consider a sentence ungrammatical for a time and then decide, upon further reflection (i.e., when more of the grammar is considered) that it is grammatical, or vice versa.

V. CONCLUSIONS

To our mind, any model of natural language which is to be useful must concern itself with the kinds of issues and phenomena we have discussed: the fluidity, imprecision, situation sensitivity, and inconstancy of language, and the robustness of human understanding. The tradition has tacitly opted for monotonic and optimizing models, not so much because of any definite arguments in favor of such models as because of a widespread failure to concern oneself with these kinds of issues in the first place. The notion that natural languages exist as ideal entities in their own right naturally leads to models that in effect are monotonic and optimizing, in the same way that considering classic logic instead of actual human reasoning or ideal economic systems instead of real ones will lead to analogous results. If we consider human linguistic behavior as our domain of study, then the satisificing and nonmonotonic models are almost inevitable. Teuchistic processes are appropriate to characterize these models.

References

Amsler, R. A. Words and Worlds. *Proceedings of TINLAP-3*. Las Cruces, NM: New Mexico State University Computing Research Laboratory, 1987, 16–19.

Charniak, E. A parser with something for everyone. In M. King (Ed.) *Parsing natural language*, New York: Academic Press, 1983, 117–149.

Chomsky, N. *Some concepts and consequences of the theory of government and binding*. Dordrecht: Foris, 1986.

Cullingford, R. E. *Natural language processing: A knowledge-engineering approach*. Totowa, NJ: Rowman & Littlefield, 1986.

Hill, A. Grammaticality. *Word, 17*, 1961, 1–10.

Itkonen, E. *Grammatical theory and metascience*. Amsterdam: John Benjamins, 1978.

Lindsay, R. K. Jigsaw Heuristics and a Language Learning Model. In N. Findler and B. Meltzer (Eds.) *Artificial intelligence and heuristic programming*. Edinburgh: Edinburgh University Press, 1971.

Lytinen, S. L. Dynamically combining syntax and semantics in natural language processing. *Proceedings of AAAI-86*, 1986, 574–578.

Manaster-Ramer, A. Copying in natural languages, context-freeness, and queue grammars. *Proceedings of Association for Computational Linguistics-86*, 1986, 85–89.

Marcus, M. *A theory of syntactic recognition for natural language*. Unpublished doctoral dissertation, MIT, 1978.

McCawley, J. D. A note on multiple negations, or, Why you don't not say no sentences like this one. *Grammar and Meaning*, 1973, 206–210.

Pāṇini. n.d. *Aṣṭādhyāyī*. Published 1887 as *Pāṇini's Grammatik* (O. Böhtlingk, ed.), Leipzig; reprinted Hildesheim: Olms, 1964.

Paul, H. *Prinzipien der Sprachgeschichte*. Halle: Niemeyer, 1880.

Peters, S. Discussion of W. C. Rounds's paper 'The relevance of computational complexity theory to natural language processing', *Proceedings of the Conference on the Processing of Linguistic Structure*. Palo Alto CA: Center for the Study of Language and Information, 1987, to appear.

Reiter, R. A logic for default reasoning. *Artificial Intelligence*, 1980, *13*, 81–132.

Riesbeck, C., and Schank, R. C. *Comprehension by computer: Expectation-based analysis of sentences in context*. Technical Report 78, Department of Computer Science. Hartford CT: Yale University, 1976.

Robinson, J. *Extending grammars to new domains*. Report RR-83-123. Los Angeles: Institute for Science and Information, 1984.

Simon, H. A. *Administrative behavior*. New York: Macmillan, 1947.

Simon, H. A. A behavioral model of rational choice. *Quarterly Journal of Economics*, 1955, *69*, 99–118.

Small, S. *Word expert parsing*. Unpublished doctoral dissertation, Department of Computer Science, University of Maryland, 1980.

Turner, R. Semantic theory of non-monotonic inference. In R. Turner (Ed.) *Logic for artificial intelligence*. Chicester: E. Horwood, New York: Halstead Press, 1984.

Wilks, Y. An intelligent analyzer and understander of English. *Communications of the Association for Computing Machinery*, 1975, *18*, 264–274.

Winograd, T. *Understanding natural language*. New York: Academic Press, 1972.

Woods, W., Kaplan, R., and Nash-Webber, B. *The Lunar Sciences Natural Language Information System: Final report*. Technical Report 2378. Cambridge MA: Bolt, Beranek, and Newman, 1972.

Yngve, V. A model and an hypothesis for language structure. *Proc. of the American Philosophical Soc.*, 1960, *104*, 444–466.