DOCUMENT RESUME

ED 070 344 FL 003 658

TITLE	Research Trends in Computational Linguistics. Conference Report.
INSTITUTION	Center for Applied Linguistics, Washington, D.C.
SPONS AGENCY	National Science Foundation, Washington, D.C. Div. of Social Sciences.
REPORT NO	NSF-GS-32399
PUB DATE	72
NOTE	98p.; Report of a Conference Sponsored by The Center for Applied Linguistics and the Association for Computational Linguistics, Washington, D.C., March 14-16, 1972
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Algorithms; Artificial Speech; *Computational Linguistics; Computer Science; *Conference Reports; Dialect Studies; Grammar; Information Processing; *Language Research; Linguistic Performance; Machine Translation; Man Machine Systems; Phonology; Psycholinguistics; *Research Needs; *Research Problems; Semantics; Sociolinguistics; Speech; Syntax

ABSTRACT

This document contains the reports summarizing the main discussion held during the March 1972 Computational Linguistics Conference. The first report, "Computational Linguistics and Linguistics," helps to establish definitions and an understanding of the scope of computational linguistics. "Integrated Computer Systems for Language" and "Computer-Oriented Grammars and Parsing" deal with technical issues and immediate concerns for research. "Machines and Speech" reports on a topic of emerging relevance in the field; lengthy consideration is given to reading machines for the blind and speech understanding systems, among other topics. "Language Performance (Psycholinguistics and Dialectology)" describes another fairly new application for computational linguistics. The last two reports, "Social Implications of Automatic Language Processing" and "Professional Ethics, Standards, and Education," involve issues related to the potential impact and implications of computational linguistics and its concerns. (VM)

RESEARCH TRENDS IN COMPUTATIONAL LINGUISTICS

Report of a Conference sponsored by
The Center for Applied Linguistics
in cooperation with
The Association for Computational Linguistics
March 14 - 16, 1972

Joyce Friedman and A. Hood Roberts, Co-Chairmen
Adam G. Woyna, Coordinator

CENTER FOR APPLIED LINGUISTICS: 1972

The Conference reported herein was supported by a grant from the National Science Foundation, Division of Social Sciences, Special Projects Program, Grant No. GS-32399.

Center for Applied Linguistics
1611 North Kent Street
Arlington, Virginia 22209

PREFACE

The Conference on Research Trends in Computational Linguistics was held in Washington, D.C., March 14-16, 1972, under the auspices of the Center for Applied Linguistics in cooperation with the Association for Computational Linguistics and with the support of the National Science Foundation. Its purpose was to assess the current situation in computational linguistics and to project future trends and needs.

The invitation list was prepared in such a way as to reflect the diversity of current work and to stretch somewhat the traditional scope of the field by including phonologists and psycholinguists, whose work will increasingly interact with computational linguistics in the years to come. The area of linguistic data processing, on the other hand, was excluded as adequately treated elsewhere.

The Conference was organized into workshop sessions; the individual groups' discussions were later summarized by the chairmen and presented in general sessions for comment and reaction. The task of summarizing the views of a workshop group into a coherent report is a difficult one; the reports benefit by being tempered by the personal views and knowledge of the particular experts chosen as chairmen. Additional comments were obtained when the preliminary versions were presented at a special session of the Tenth Annual Meeting of the Association for Computational Linguistics at Chapel Hill, July 1972.

"Computational Linguistics and Linguistics", the first plenary session, set the tone of the meeting; as Barnes' report shows, some agreement evolved on a definition and understanding of the scope of computational linguistics. "Integrated Computer Systems for Language" were also discussed in a plenary session. Its chairman, Simmons, summarizes the discussion, adds a personal statement of his view of the field, and appends an outline of suggestions for future research.

Three substantive subject areas were discussed at length by smaller groups. Almost all of the participants in Petrick's sessions on "Computer-Oriented Grammars and Parsing" have at one time or another been engaged in producing such grammars. His report includes first-hand summaries of current work. As Cooper notes, "Machines and Speech" was on the agenda not because of past performance but because of future expectations. Reading machines for the blind and speech understanding systems form the initial points of contact; Cooper spells out progress to date and lists research needs. "Language Performance" constitutes another topic of potential interest to the field. Carroll surveys the past applications and concludes that present work in computational linguistics can benefit psychological and psycholinguistic studies.

Important non-technical questions were also discussed. Walker's essay on "Social Implications of Automatic Language Processing" reflects concern lest computational linguists' work be co-opted for anti-social purposes; overpowering this concern there is stress on the positive implications of the field. Sedelow's session treated the mixed topics of "Professional Ethics, Standards, and Education".

The preceding paragraphs present a brief outline of the content of the meeting and the resulting reports. It is important to add that although the participants in the meeting expressed many diverse views, there was a pervading feeling of common interest and optimism for the future of computational linguistics. The research problems facing the field are difficult but immensely interesting; recent developments leave computational linguists hopeful.

PARTICIPANTS

Robert F. Barnes, Jr.
Lehigh University

John B. Carroll
Educational Testing Service

Franklin S. Cooper
Haskins Laboratories

Bozena Dostert
California Institute of Technology

Harold P. Edmundson
University of Maryland

Joyce Friedman, Co-Chairman
University of Michigan

Paul L. Garvin
State University of New York at Buffalo

Joseph E. Grimes
Cornell University

David G. Hays
State University of New York at Buffalo

Aravind K. Joshi
The University of Pennsylvania

Ronald M. Kaplan
Harvard University

Martin Kay
RAND Corporation

Dennis Klatt
Massachusetts Institute of Technology

Sheldon Klein
The University of Wisconsin

William N. Locke
Massachusetts Institute of Technology

Ignatius G. Mattingly
The University of Connecticut

John A. Moyne
Queens College

Edward P. Neuburg
National Security Agency

Milos Pacak
National Institutes of Health

Stanley R. Petrick
IBM Corporation

Arnold W. Pratt
National Institutes of Health

A. Hood Roberts, Co-Chairman
Center for Applied Linguistics

Jane J. Robinson
IBM Corporation

Gerald M. Rubin
Brown University

Naomi Sager
New York University

Sally Y. Sedelow
The University of Kansas

Robert F. Simmons
The University of Texas

Bernard Vauquois
The University of Grenoble

Donald E. Walker
Stanford Research Institute

Terry Winograd
Massachusetts Institute of Technology

William A. Woods
Bolt, Beranek and Newman Inc.

Adam G. Woyna, Coordinator
Center for Applied Linguistics

Victor H. Yngve
The University of Chicago

Observers

Murray Aborn
National Science Foundation

Peter G. Lykos
National Science Foundation

Gordon R. Sherman
National Science Foundation

Val G. Tareski
National Science Foundation

TABLE OF CONTENTS

COMPUTATIONAL LINGUISTICS AND LINGUISTICS

Robert F. Barnes, Jr., Chairman

As various participants began to express their views on the nature and status of Computational Linguistics (CL) and on its relation to Linguistics in general, it seemed at first that a consensus might be a hopeless goal. But as the discussion continued, two things gradually became clear: (1) that CL does indeed comprise a broad range of studies, activities, and techniques; and (2) that, as a result, apparent disagreement between differing views on CL is often simply a reflection of differences in relative emphasis among these areas.

For example, to computational linguists concerned chiefly with computer implementation of natural-language grammars, CL seems largely to be an "experimental branch" of its parent field. But further discussion brought out clearly that CL is far from being merely the linguistic counterpart of, let us say, experimental physics; a number of genuinely theoretical areas and questions exist in CL, too ---theorems relating formal grammars to automata are but one example. Other speakers suggested an opposite extreme---that the entire field of linguistics itself (at least insofar as its empirical aspects are concerned) is essentially computational in its reliance on the computer. From this point of view, there is no distinction at all between the two fields, and the very term "computational linguistics" seems distinctly odd---as if we spoke of "microscope biology" or "slide-rule engineering".

But in some areas of linguistics (especially the more traditional ones), computers are used as convenient tools, which in principle, could be replaced with an army of slaves (or perhaps graduate assistants). In other areas, however, computational notions play a more conceptually central role. Thus, without completely excluding from CL such activities as the development and application of computer procedures to perform traditional linguistic activities, we can distinguish a narrower core area in which notions from mathematics and computation interact crucially with those from linguistics. The central focus of this interdisciplinary area is the linguistic (or more generally, symbolic) algorithm, and it is with the structure and application of these that CL is peculiarly concerned.

It should be noted that this concern makes CL a genuine interdisciplinary field in its own right, not merely a sub-field of either linguistics or computation. Indeed, some of the particular concerns of CL may go considerably beyond what is customarily seen as lying in either of these spheres of interest. A computational linguist working on problems in compiler syntax and semantics, for example, might well employ concepts, techniques, or results typical of CL without being seen as doing work in either linguistics or mathematics. Another computational linguist (or indeed, even the same person) might also be working on algorithms directed toward the confirmation or disconfirmation of some purely linguistic theory; such an activity may be quite machine-free and may not at all be seen as belonging to the field of computation---either in theory or practice. Furthermore, certain applications of CL techniques, such as automatic style- or content-analysis of natural-language source material, belong even less uniquely either in linguistics or in computation. Still more remote from both fields are such activities as the application of generative grammars to musical composition---that is, natural extensions of certain indisputably computational-linguistic studies which may well be seen as constituting peripheral areas of CL. But throughout this range of activities, the essentially algorithmic nature of CL gives it a thoroughgoing air of rigor which is not always present in the parent field of linguistics; and its central concern with the symbolic nature of its source material distinguishes it from the general field of computation.

It should also be noted here that though almost everyone feels some degree of uneasiness with the term "Computational Linguistics" itself, no one seems able to offer a clearly superior substitute. Broader terms have been suggested but they seem no less vague and indefinite, and sometimes are so broad that it would be hard to say what is not CL. It may well be that for our purposes a vague, ill-defined, but relatively narrow term is better than a vague, ill-defined broad one. And in any case, as a field thrives and grows, the origins of its title are of considerably less importance than its identity as a healthy intellectual discipline.

As the group's discussion continued, there seemed, therefore, to come into focus (in addition to the sort of computational assistance utilized by traditional linguistic studies) two characteristic sorts of activities that perhaps implicitly define CL---(1) the application of linguistic concepts and techniques to problems dealing with computers and computation, and (2) the algorithmic interpretation or realization of linguistic problems or theories. (Sometimes, of course, as in the development of grammar-testing routines, a single activity may be seen as belonging to both.) In each of these areas, but especially in the latter, CL takes on a special character as a discipline---neither fully theoretical nor fully empirical, but serving as a rigorous "bridge theory" connecting other, more abstract theories with the specific facts they are to account for. It is in just such a role, for example, that studies in CL may prove helpful in connecting the notions of linguistic "competence" and "performance" in an enlightening, yet formally precise way.

It has been generally accepted since de Saussure that linguistics need not concern itself only with describing actual utterances, with their idiosyncratic ranges of linguistic features, many of them "accidental". But exactly which features are seen as accidental and which as essential depends very greatly upon one's underlying linguistic "paradigm" (to use Thomas Kuhn's helpful concept). A key issue between the post-Chomskian and pre-Chomskian paradigms for linguistics centers on the role of performance in a linguistic theory. The one sees performance as empirical data---evidence for or against an abstract theory of competence; the other sees it as an empirical phenomenon---the very subject matter to be accounted for by the theory. (John Searle, in the New York Review of June 29, 1972, puts this distinction very nicely.)

Now, when theory and data fail to match, it need not be the theory which is rejected; if one can provide a systematic explanation of the mismatch, the theory and data can still be seen as compatible. The abstract theories of the transformationalists (as well as the transformationalists themselves) are often criticized on their seeming disregard of facts of linguistic usage---the performances we encounter every day. The transformationalists generally respond that accidental features of performance have created apparent inconsistencies, which can ultimately be explained at some later point in the theoretical development. Who has the better part of the argument? In the absence of some systematic way to relate theoretical competence to empirical performance, it's hard to say. It seems possible, though, that the rigorous nature of CL, structured as it is about algorithms (which can be seen as idealized performances of a certain sort), may offer exactly this needed connection between competence and performance.

As one clarifies the nature of CL, one almost inevitably explores what its future prospects may be, as shown in the preceding remarks. In this direction, a brief glance at the history of the field also reveals two significant facts which suggest, perhaps, where CL's most fruitful development may lie. First, over the past years, one can note a definite trend away from narrow studies oriented solely toward particular applicational goals. In part, this represents the maturing of the discipline as its own characteristic problems emerge and increasing attention is directed to them. However, there seems to be an additional feature whose significance is emphasized by the second point---namely, that these earlier "mission-oriented" studies were often conducted in a framework of (then-existent) theory, often brought in from outside the discipline, which has since been recognized as inadequate. (What this says in general about the prospective adequacy of specific, mission-oriented theory may be left to speculation.) More recently, however, one can identify a rich stage of theory arising from within CL itself. The theories of this period are often considerably broader than the older ones; they often span the entire syntactic-semantic-pragmatic spectrum.

Some of this new theory in CL appears to be relatively "conventional" in the sense that it represents new advances along previously established theoretical lines; new parsing procedures for context-free grammars would perhaps serve as examples of developments of this sort. Many of the new developments, however---particularly in areas associated with semantic information processing---have involved the formulation of their own theoretical structures. For example, a long-standing framework for the study of language, widely accepted in logic and the philosophy of language, rests on a threefold distinction between syntax, semantics, and pragmatics---roughly, the theories of form, meaning, and use. Chomsky' early arguments against basing grammatical theory on meaning, as well as the then-prevalent structuralist attitude toward meaning, were in part based on this distinction. More recently, however, there are indications that an adequate theory of syntax may require semantic and pragmatic components of some sort, but the proper theoretical organization of such components is still very uncertain. In CL, on the other hand, some of the most successful work in semantic information processing has tended not to presuppose the syntactic-semantic-pragmatic distinction, or else to utilize other distinctions which might be vaguely similar but which are made on somewhat different grounds. In these and a number of similar instances, CL has begun to show an intellectual maturity and self-sufficiency which is genuinely promising.

As a result of this increased theoretical development, one may expect CL not only to concern itself with its "own" questions and problems, but also to contribute toward broader areas in linguistics it has tended to neglect in the past. Computational implementation of phonological theories---in either generative or analytical programs---may provide fruitful contributions to work on speech synthesis and recognition, as well as valuable empirical test procedures for more general phonological theories. Developments in computational semantics (such as those previously mentioned) may offer mechanisms for an apparently needed greater theoretical richness and structure in semantics, as well as for applications in grammar and phonology. Additionally, by presenting its own new viewpoints toward such orthodox linguistic distinctions as performance vs. competence, declarative vs. imperative sentence, syntax vs. semantics vs. pragmatics, etc., work in CL may contribute considerably to the clarification of the foundations of linguistics.

(Though they weren't mentioned explicitly in the meeting, David Hays's two recent papers on the nature of CL*, based on remarks made at the 1971 International Meeting in Hungary, are so to the point that they must be mentioned here. While the viewpoint presented there is not precisely that given here, the two seem more complementary than contradictory; and the papers should be read by all concerned with the future of the field.)

In sum, CL is not linguistics---in whole or in part. Its concern is legitimately far broader than natural language, though much of its attention is directed toward problems involving automatic processing of natural language material. Its central focus---the notion of symbolic algorithm---represents a standard in rigor and precision which general linguistic procedures do not always attain. On the other hand it is quite clear that CL does have close historical, conceptual, and practical ties to linguistics; and although we expect CL to develop in such a way as to make more precise its own characteristic nature, we also expect the relationship between the two fields to remain close and mutually profitable.

---

*"The Field and Scope of Computational Linguistics" and "The Past, Present, and Future of Computational Linguistics", available from the author, Department of Linguistics, State University of New York at Buffalo.

INTEGRATED COMPUTER SYSTEMS FOR LANGUAGE

Robert F. Simmons, Chairman

Most recent and notable among integrated systems for language processing are Woods' Natural Language Retrieval System and Winograd's system for understanding commands to a robot for operations on a limited world of colored blocks. Other examples of more or less successful integrated language processing systems include mechanical translation, natural language data manipulation, robot command systems, natural language CAI, semantic analysis of medical data, the analysis, synthesis and testing of linguistic rules, text analysis, paraphrase and question answering, and text generation. The current Advanced Research Projects Agency (ARPA)-sponsored research toward useful speech recognition systems is the most recent large scale approach to integrate knowledge of semantics, syntax and phonology into a capability for computer understanding of a useful spoken vocabulary. From this inventory of examples we can infer that an integrated language processing system is a processor that organizes a set of component linguistic processes into an effective and potentially practical language processing device. The components may include speech analysis and synthesis, syntactic, semantic and logical operations.

One aspect of integrated language processing can arbitrarily be dimensionalized as the vertical inter-relation of morphological, syntactic, semantic and pragmatic information. On this dimension, interaction among levels is used to resolve ambiguities and to clarify understanding. The most thorough example currently available as a working program appears to be Winograd's system that uses syntactic, semantic, and pragmatic information to resolve possible ambiguities of English commands and questions. The complexity of integrated systems in their vertical dimension is such that a single person can hardly hold all aspects of their components' interaction in his mind at one time. This level of complexity requires the use of high level languages such as LISP, PLANNER, etc., with the consequence that computation is fairly slow and fast random access memory requirements are quite large. Alternatively, such a system may be built as a team effort, with its attendant communication problems among team members and subprograms. The requirement for careful documentation of such systems was emphasized and suggestions emerged for the use and development of text editing and automated documentation-program packages. Automated flow-charting programs were suggested for application to FORTRAN, ALGOL, and PL/I type programs, and systems for displaying the calling structure of functions and other organizational structures of LISP systems were briefly described.

A second aspect of integrated systems can arbitrarily be dimensionalized as horizontal. On this dimension problems occur at the interface of a given language processing system with the user and with other systems. Careful human engineering of the user interface is required to enable the convenient insertion of lexical entries, grammar rules and semantic and pragmatic information as well as the testing of these data for consistency and accuracy. The need for fast on-line interactive consoles is most obvious here. Problems concerned with evaluation of the effectiveness of language processing systems and their generalizability emerge when considering horizontal integration. These areas are suggested for continued research.

A contrast emerged between what is currently desirable and what is minimally necessary in the way of hardware for research and development of integrated language processing systems. Ideally, something like the following configuration of hardware is desired:

INPUT	CENTRAL PROCESSOR	OUTPUT
Character Reader Remote Quiet Terminals Voice Input Graphics	Core Memory ¼-1 million* words Extended Core 1 million words Disc 10 million words Networked Miniprocessors	Print Voice Graphics

*numbers signify order of magnitude only

Software support should include a multi-access timeshared system, basic languages such as FORTRAN, ALGOL, PL/I, LISP, and SNOBOL. Specialized text editing, parsing, generation and question answering program systems would serve as a computational library for the system.

In constrast to the ideal system designed to support integrated language processors, the minimum requirements can already be found at a number of locations around the country in existing computing and communication networks. The Massachusetts Institute of Technology AI laboratory; Bolt,Beranek & Newman Inc.; Stanford University; Carnegie-Mellon University; University of Texas; Stanford Research Institute; and the ARPA computing research network are among a number of computing facilities adequate to the purpose. Unfortunately, adequate access to such facilities is administratively limited and very expensive.

A significant question was raised about hardware requirements which led the discussants to conclude that present day hardware systems are theoretically adequate to support research in computational linguistics but, in fact, the research user typically commands less computer utility in terms of central processors and memory than he did in the mid-sixties. The reason is that despite the truly impressive increases in size of memory and processor efficiency, the computing utility is shared much more widely, usually reducing the amount of utility that each person can command. The consequence appears to be that computational linguists often find it difficult to obtain sufficient computing power to construct large integrated systems. For the moment, this lack does not block research; but it does bias the activity toward the development of component software rather than large integrated systems which require more computation utility than is readily available to the researcher.

The suggestion was made that computational linguistics research might follow the path customarily used by atomic physicists in concentrating expensive equipment at a few centers. The ARPA network is a good example of this approach in the related field of artificial intelligence research, and it might well be recommended for our own attempts to develop practically useful integrated language processing systems.

The foregoing paragraphs represent the essence of the morning's discussion on Integrated Systems in computational linguistics. Behind this discussion lay a considerable amount of common knowledge of the current state of the art, the research problems and the implications of integrated systems for the future of the discipline. In this section an attempt is made to render some of this implicit background explicit to the reader. The effort is of course significantly biased by the author's perception of the situation.

First it appears that there has been significant development of syntactic analysis procedures over the past few years. Petrick's report outlines the state of the art in terms of the many parsers that are currently more or less well-developed at various locations. One may note in particular that many of the mature projects such as those reported from Europe, the still-healthy mechanical translation project of Lehmann and Stachowitz (1972), and numerous newer language processing systems around the country, including those of Winograd (1972), Thompson (1964), Carbonnel (1970), Heidorn (1971), and Simmons (1972), did not pattern themselves on the strict transformational model. Some systems such as those of Woods (1970) and Kellogg (1968), while originally based on transformational theory, reveal their transformational origin mainly by their use of deep syntactic structures and some use of selection restrictions and semantic markers.

Placing these observations in the context of: (1) the rapid development of case structure theory, (2) Winograd's application of

Halliday's (1970) syntactic approach, and (3) the computational development by Woods (1970) and by Bobrow and Fraser (1969) of the Thorne, Bratley and Dewar (1968) use of a finite state network approach to grammar, it is possible to see the emergence of a new theoretical framework for computational linguistics. This approach is somewhat eclectically based on the use of some form of procedural grammar that is represented as rule sets (see Heidorn), as augmented finite state nets (see Woods), or as procedures (see Winograd, Schank).

For structures to represent the underlying meaning of sentences, three approaches are quite prominent. To the extent that a phrase refers to some object in a data structure, it is quite valuable to follow the procedural semantics approach described most thoroughly by Woods (1968). In this approach the meaning of a phrase is the value returned by a procedure that uses elements of the phrase to identify a data object or set as its referent.

Winograd carried the procedural semantics approach even further by interpreting relational words such as verbs and prepositions as procedures in a robot command language. The MicroPlanner Interpreter of this language is sophisticated enough to accomplish quite significant inferences. Consider, for example, the command, "Put the red block in the blue box", which translates into a procedural expression such as, (PUT B1 B2). If the blue box, B2, already contains an object, the PUT function is able to call other functions to temporarily release the block, B1, remove the object from B2, then pick up B1 again and finally succeed in putting it in the box.

A second approach that is theoretically satisfying but computationally cumbersome is to represent the meaning of a sentence as a statement in symbolic logic. Such a statement is then taken as a theorem whose truth value with reference to a set of axioms that comprise the data base can be determined by use of resolution techniques. (See Green & Raphael, 1968; Palme, 1971; and Sandewall, 1970).

A third approach---one that is applicable to text analysis---is to resolve the sentence into a deep-case representation which is conceptual ized as an event description in terms of a verbal relation and its arguments. The meaning of this event description is a series of propositions implied by the event. (See Schank, Lehmann and Stachowitz, and Simmons.)

Since the third approach is rather new, it would best be illustrated by the following example:

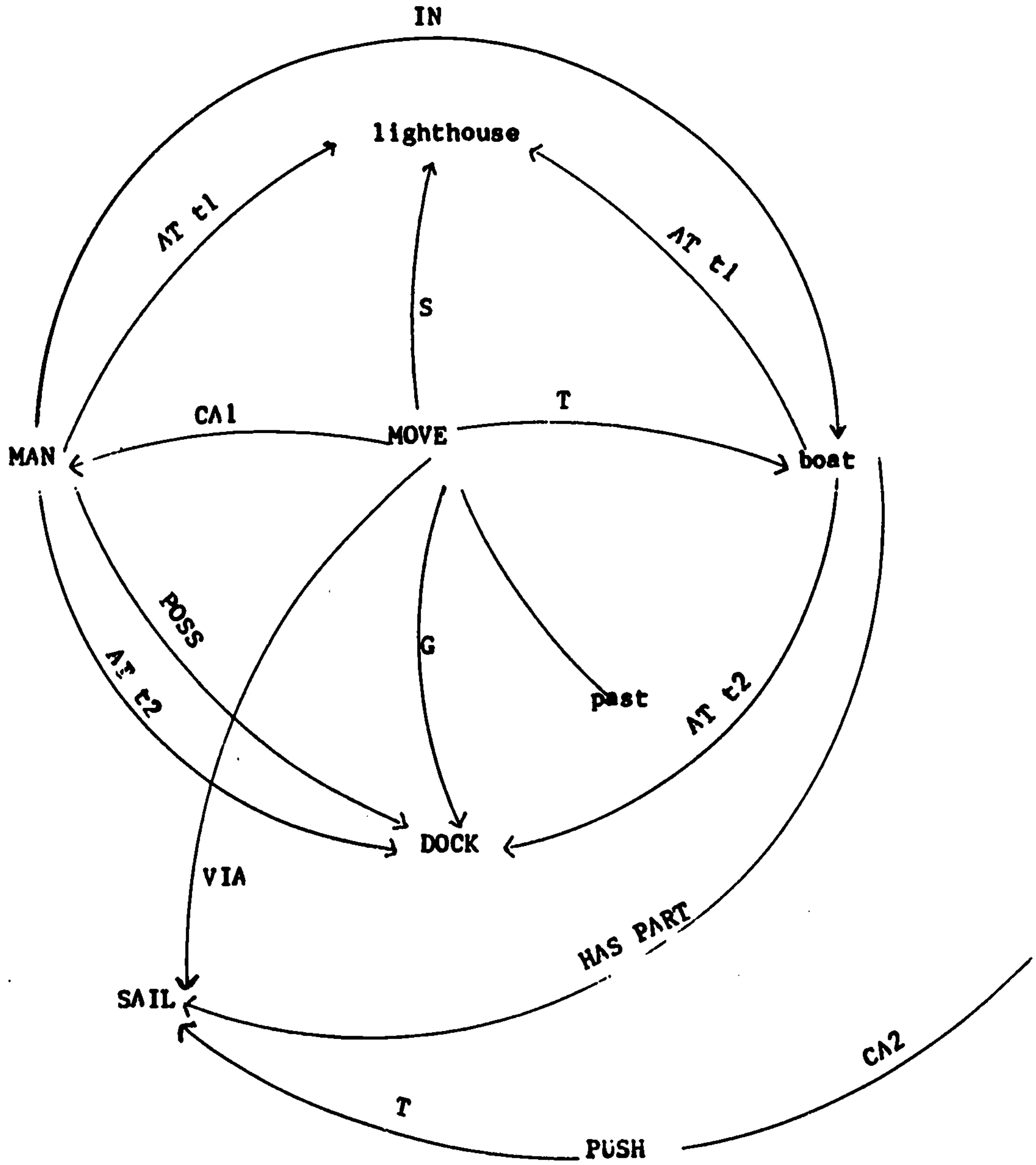
The man sailed a boat from the lighthouse to his dock.

A form of its deep case structure is as follows:

MOVE	via sail
CAUSAL ACTANT 1	the man
THEME	boat
SOURCE	the lighthouse
GOAL	his dock

Some of the propositions implied by this structure can be seen in the diagram of Figure 1. Each node-arc-node in that diagram represents such a proposition. For example, "MAN Causal Actant of MOVE" shows that the man instigated the action; "MOVE THEME BOAT" shows the boat received the action; "MAN Possess DOCK" represents the meaning of "his dock"; "BOAT At, t1 LIGHTHOUSE" shows an initial state of the event; and "MAN At, t2 DOCK" shows that the position of the man at the end of the event was that of the dock. The other arcs can be interpreted in similar fashion. The entire set of such arcs provides a first level of inferential meaning for such purposes as question answering and problem solving.

These three approaches to representing the meaning of English sentences are quite obviously the current computational nominations for semantic deep structures underlying the surface form of sentences. Each has proved useful in some integrated system application. There is



Graph of Sentence Meaning

Figure 1

enough similarity among the approaches to suggest that they may in fact be different points of view and that a given semantic structure has procedural, logical and propositional descriptions.

The means for computing these underlying structures are remarkably similar at a certain level of abstraction. The surface structure of an English sentence is analyzed with the aid of a grammar, a lexicon and a parsing algorithm to produce noun phrase and verb phrase components. Either at the point of recognizing a phrase, or later, a series of operations is applied to transform each phrase into the desired form of semantic structure. Although syntactic analysis remains a difficult, indeterminate, recursive process, the clear definition of the desired semantic structure is used successfully to simplify the process. The validity of the resulting semantic structure is measured by its usefulness in question answering, paraphrase, and inferential capabilities. Rationalist arguments about the exact form and time of application of transformational rules are essentially irrelevant in this empirical approach, and the definition of semantic deep structure is influenced most strongly by pragmatic considerations.

What we see in looking briefly at the development of integrated systems for understanding English is that computational linguistics is scientifically past its adolescence. While not being "school-bound", it is nevertheless deeply influenced by such syntactic theorists as Harris, Chomsky, Halliday, Thorne and others. It reflects semantic theories of Tarski, Carnap, Katz, Woods, and most recently Fillmore, Chafe and Grimes. Its methods clearly demonstrate the usefulness of computational theories of Thompson, Kuno and Oettinger, Woods and others. The field is also fortunate in that a group of psychologists are exploring the applicability of its case models and procedural semantic structures as models of human memory processes. (See Frijda, 1972; Collins and Quillian, 1969; Norman, 1972; and Frederikson, 1972; among others.) And Carroll, in his report, suggests other valuable relations between computational linguistics and psychology.

But post-adolescence is not maturity. Computational realizations of natural language processing capabilities are impressive in their ability to understand small subsets of natural language but a great deal of development is required to achieve practical applications.

Most existing systems are dealing only with single sentences and only the most recent few resolve some simple cases of pronominal and anaphoric reference. Wilks (1968), Heidorn (1971), and Su (1971) have studied paragraphs or larger units of text, but these appear to be initial explorations of a very large area of computational linguistic research. Few of the systems have utilized peripheral storage to allow for growth of the large lexicons and grammars that are usually required for practical applications. The task of producing large, consistent lexicons and grammars has only occasionally been attempted. (See Woods, 1970; Sager, 1970; and Lehmann and Stachowitz for examples of such attempts.)

The capability to analyze unedited text---essential for many mechanical translation and information retrieval applications---is far out of reach at present, depending on the development of large computational lexicons and grammars for a large stable system with much use of peripheral storage. An area in which very little work has so far been accomplished is that of developing convenient means for nonlinguists to communicate lexical and grammar structures to the system.

But applications of this post-adolescent field of computational linguistics constitute the most active area of research. Early work on stylistic and content analysis (see Sedelow, 1969; Stone, 1962) has resulted in tools frequently used by behavioral scientists and humanities scholars for a wide variety of purposes. Another early development of keyword analysis of sentences is still a major tool for information retrieval applications. Several efforts at natural language data management (see Woods, Thompson and Kellogg for examples) have explicitly practical goals. A most recent applications area is that of Computer Assisted Instruction with studies by Wexler (1970), Carbonnel (1970), Brown (1972), Simmons (1972), and others. The attempt in this area is to use language processing and question answering technology to develop some approximation to a tutorial system that can understand a student's questions and statements, and his answers to the system's questions. The systems available at this writing are purely experimental---sometimes impressive in the depth of their understanding of a few concepts---but lacking both breadth of language processing ability and depth of peripheral memory-system engineering. Even more important, almost nothing has so far been learned on an experimental basis of how to use a language processor as part of an effective tutorial system. A pressing need is upon us to substitute experimental evidence for armchair theorizing about the educational use of natural language tutorial systems.

Recommended Research Emphases

The foregoing outlines the area of integrated systems research as central to the discipline of computational linguistics which is based on syntactic and semantic theories of classical linguists and logicians and computational theories of its own. It has been described as an empirically oriented area with psychological implications for understanding human thought processes and one with potentially great---but still unrealized---applications to socially valuable systems. The meeting's discussion on integrated systems was rich in its suggestions for research and in its discussion of current difficulties in accomplishing it. Other areas requiring research become apparent in considering the field's state of the art.

An attempt to summarize the group's suggestions in outline form is presented below. Each of the subtopics represents an area where research or developmental attention can be expected to advance the state of the art.

I. Managing complex systems
   1. High level languages
   2. Team efforts
      a) Communications among members
      b) Communications among system components

3. Documentation and programming aids
   a) Text and file editing
   b) Automated flow-charting systems
   c) Displays of structure of subsystems

II. Interfaces
   1. User
      a) Convenient methods for inserting data
      b) Testing, revising data
      c) Evaluation of language processor (including choice of standard dimensions for evaluation)
   2. Peripheral systems
      a) Speech analysis
      b) Speech synthesis
      c) Peripheral memory devices
      d) Text scanners

III. Computing utility
   1. Central repository of systems, data, algorithms
   2. Distant and local access to a very large computing system (or network)

IV. Theoretical research
   1. Syntactic structures for computation
   2. Semantics of case structures, procedures, etc.
   3. Psychological studies of memory structures
   4. Logical structure of sentence meanings
   5. Speech

V. Computational linguistic research
   1. Discourse structure
   2. Anaphoric reference
   3. Computational structure of the English lexicon
   4. Development of computational lexicon and grammar for large subsets of English (and other languages)
   5. Computational structure of speech

VI. Systems research

1. Large language processing systems with indefinitely extendable peripheral storage capabilities
2. List processing languages that can effectively use peripheral storage devices

VII. Applied natural language systems

1. Computer assisted instruction
2. Text-based question answering
3. Data-based question answering
4. Natural language systems for describing algorithms and processes to a computer, i.e., Natural Language Programming Systems
5. Systems for conversing with a user about a topic
6. Systems for generating essays from a data base
7. Simulations of personality and belief structures

VIII. Inferential capabilities

1. Verbal problem solving
2. Theorem proving
3. Deductive question-answering
4. Paraphrase

IX. Esoterica

1. Emotional content of language and of speech
2. Interface with visual data
3. Interface with auditory data
4. Stylistics, rhythm, rhyme, etc.
5. Semantic classification systems

References

Bobrow, D.G., and B. Fraser. "An Augmented State Transition Necwork Analysis Procedure." In: Proceedings of the International Joint Conference on Artificial Intelligence, 557-567. Washington, D.C., 1969.

Brown, John S., R.B. Burton, and Frank Zydbel. A Model Driven Question-Answering System for a CAI Environment. (Technical Report #13.) Irvine, Calif.: Department of Information Science, University of California, Irvine, January 1972.

Carbonell, J.R. "AI in CAI: An Artificial Intelligence Approach to Computer Assisted Instruction." IEEE Transactions on Man-Machine Systems MSS-11(1970):180-202.

Carnap, R. Introduction to Semantics. Cambridge, Mass.: Harvard University Press, 1946.

Chafe, W.L. Meaning and the Structure of Language. Chicago: University of Chicago Press, 1970.

Collins, A.M., and M.R. Quillian. "Retrieval Times from Semantic Memory." Journal of Verbal Learning and Verbal Behavior 8(1969):240-247.

Fillmore, C.J. "The Case for Case." In: Emmon Bach and Robert T. Harms, eds. Universals in Linguistic Theory, 1-88. New York: Holt, Rinehart and Winston, 1968.

Frederiksen, Carl H. Representing Logico-Semantic Features of Written Knowledge Acquired from a Discourse. Berkeley, Calif.: Department of Psychology, University of California Berkeley, 1972 (author ms).

Frijda, N.H. "The Simulation of Human Memory." Psychological Bulletin (January 1972).

Green, Cordell, and B. Raphael. "The Use of Theorem Proving Techniques in Ouestion Answering Systems." In: Proceedings of the 23rd ACM National Conference, 169-181. New York, 1968.

Grimes, Joseph E. The Thread of Discourse. (Technical Report No.1, National Science Foundation Grant GS-3180.) Ithaca, N.Y.: Department of Modern Languages and Linguistics, Cornell University, 1972.

Halliday, M.A.K. "Functional Diversity in Language as Seen from a Consideration of Modality and Mood." Foundations of Language 6(1970):322-361.

Heidorn, George E. Natural Language Inputs to a Simulation Programming System. Monterey, Calif.: Naval Postgraduate School, December 1971.

Kellogg, C.H. "A Natural Language Compiler for On-Line Data Management." In: AFIPS Conference Proceedings. Vol. 33, Part 1. 1968 Fall Joint Computer Conference, 473-492. Washington, D.C.: Thompson Book Co., 1968.

Lehmann, Winifred P., and Rolf Stachowitz, eds. Feasibility Study on Fully Automatic High Quality Translation. Final Report, 1 February 1970--30 June 1971. (Report No. RADC-TR-71-295.) Griffiss Air Force Base, N.Y.: Rome Air Development Center (IRDT), December 1971.

Norman, Donald A. Memory, Knowledge and the Answering of Questions. La Jolla, Calif.: Center for Human Information Processing, University of California, San Diego, May 1972.

Palme, J. "Making Computers Understand Natural Language." In: N. Findler, and B. Meltzer, eds. Artificial Intelligence and Heuristic Programming, 199-244. Edinburgh: Edinburgh University Press, 1971.

Sager, Naomi. An Application of Syntactic Analysis to Information Retrieval. (String Program Report No. 6.) New York: Linguistic String Program, New York University, April 1970.

Sandewall, E.J. Representing Natural Language Information in the Predicate Calculus. (CS-166, AIM-128.) Stanford, Calif.: Artificial Intelligence Project, Department of Computer Science, Stanford University, July 1970.

Schank, Roger. "Identification of Conceptualizations Underlying Natural Language." In: Roger Schank, and K. Colby, eds. Computer Cognition. San Francisco: W.H. Freeman & Co. (in press).

Sedelow, S., and W. Sedelow. "Stylistic Analysis." In: H. Borko, ed. Automated Language Processing, 181-213. New York: John Wiley, 1967.

Simmons, R.F. "Linguistic Analysis of Constructed Student Responses in CAI." In: Wayne H. Holtzman, ed. Computer-Assisted Instruction, Testing and Guidance. New York: Harper-Row, 1970.

Simmons, R.F. Semantics Networks: Their Computation and Use for Understanding English Sentences. Austin, Texas: Department of Computer Sciences, University of Texas, Austin, May 1972.

Stone, P.J., F.F. Bayles, J.Z. Namenwirth, and D.M. Ogilvie. "The General Inquirer: A Computer System for Content Analysis and Retrieval Based on the Sentence as a Unit of Information." Behavioral Science 7,iv(1962):1-15.

Su, Stanley Y.W. A Computational Model of Paragraph Production. (Technical Report No. 71-102.) Gainsville, Fla.: Center for Information Research, University of Florida, November 1971.

Tarski, A. "The Semantic Conception of Truth." Philosophy and Phenomenological Research 4(1944). [Reprinted in: Herbert Feigl and Wilfrid Sellars, eds. Readings in Philosophical Analysis. New York: Appleton-Century Crofts, 1949.]

Thompson, F.B. Semantic Counterpart of Formal Grammars. Santa Barbara, Calif.: TEMPO General Electric Company.

Thompson, F.B., et al. DEACON Breadboard Summary. (RM64TMP-9.) Santa Barbara, Calif.: TEMPO General Electric Company, March 1964.

Thorne, J., P. Bratley, and B. Dewar. "The Syntactic Analysis of English by Machine." In: D. Michie, ed. Machine Intelligence 3, 281-309. New York: American Elsevier, 1968.

Wexler, J.D. A Generative Teaching System that Uses Information Nets and Skeleton Patterns. Unpublished Ph.D. dissertation, Computer Sciences Department, University of Wisconsin, Madison, 1970.

Wilks, Yorick. Computable Semantic Derivations. (SP-3017.) Santa Monica, Calif.: System Development Corporation, January 1968.

Winograd, T. "A Program for Understanding Natural Language." Cognitive Psychology 3(1972):1-191.

Woods, W. "Procedural Semantics for a Question-Answering Machine." AFIPS Conference Proceedings. Vol. 33, Part 1. 1968 Fall Joint Computer Conference, 457-471. Washington, D.C.: Thompson Book Co., 1968.

Woods, W. "Transition Network Grammars for Natural Language Analysis." Communications of the ACM 13(1970):591-606.

COMPUTER-ORIENTED GRAMMARS AND PARSING

S.R. Petrick, Chairman

General Considerations

The first of the three sessions devoted to this topic began with an attempt to determine what is implied by the phrase "computer-oriented grammars". One interpretation suggested by some participants was that computer-oriented means "well-specified". One can utilize a grammar in a computer system only if one has stated explicitly how that grammar designates certain strings as sentences and assigns structure(s) to them. The complete specification of a grammar or a class of grammars is not directly tied to computers in any fashion. Nevertheless, an indirect association is often made for at least two reasons. First, computational linguists have of necessity been obliged to make vague or unspecified aspects of a class of grammars precise even in cases where a linguistically well-motivated defense of their choices was not attempted. Such an approach is useful in that it brings to linguists' attention problems and required choices that they may have overlooked. Second, in those cases where a linguist actually writes a non-trivial set of rules which specify a self-contained subset of some language, the use of a computer is required to verify that his grammar does what was intended. This means not only checking that representative sentences are indeed generated and assigned desired structures, but also verifying that ungrammatical sentences are excluded and that unwanted structures are not assigned. A computer is absolutely essential for accomplishing this task. Writing a set of linguistic rules, i.e., a grammar, is very much like writing a more conventional computer program. Just as in the case of a computer program, it is virtually certain that an explicit grammar, as first written, contains errors, and those errors can be removed only through a process of extensive "debugging". It is also likely that along with a host of trivial, uninteresting errors, the debugging process will also turn up some more serious, fundamental conceptual errors.

A second possible interpretation of the term "computer-oriented grammar" that was suggested was "amenable to some task related to computers". We have already accounted for the case in which the task is related to the needs of a linguist to explore the adequacy of his grammars. The task can also be one arising from an application outside of linguistics. Examples include natural language question answering systems, programming systems, computer aided instruction systems, and information retrieval systems. There is, of course, no necessity that the task be so specific. "Computer-oriented" could simply be taken to imply amenability to such a function as parsing or semantic interpretation.

The participants of the first session saw no reason to agree upon a choice between these or other possible interpretations of the phrase, "computer-oriented grammars". Instead, there was general agreement on an extensional meaning of the phrase, defined by a set of individual efforts.

Before proceeding to a discussion of those efforts, however, an attempt was made to provide a taxonomy which might be useful in grouping the grammars for discussion and in exploring their various properties. The following types of grammars were cited and very briefly discussed: context-free grammars and weakly equivalent grammars, such as categorial grammars and dependency grammars; context-sensitive grammars; general rewriting system grammars; transformational grammars based on either underlying strings or trees; augmented transition network grammars and procedural grammars; tree automata; and stratificational grammars. These classes of grammars were discussed in terms of such attributes as the adequacy of the structural descriptions they assign and the coverage of English or some other natural language which has been attained through their use. It was pointed out that coverage and adequacy of assigned structures are not independent. Some of the most successful grammars in terms of sheer coverage (e.g., context-free grammars) have been least successful in assigning acceptable structural descriptions. It was also pointed out that adequacy of assigned structures is a matter of considerable controversy in linguistics. Linguists can debate at great length whether a given structure is an adequate representation of a sentence; however we cannot begin to consider this issue meaningfully until we have specified the semantic component which is to interpret the structure in question.

In the context of such applications as machine translation, question answering systems, natural language programming systems, and information retrieval, there is much less uncertainty about whether a given sentence is treated in an acceptable manner. One would thus expect the capabilities of these systems to be well-documented. However, the facts do not bear this out. It is, in general, impossible to assess from published accounts of natural language processing systems the limits of the subsets of English which are currently treated satisfactorily by their syntactic and semantic components. Syntactic coverage is particularly difficult to determine. It appears unreasonable, to be sure, to require the writer of a formal grammar to specify the language of his grammar in an alternate, easier to understand, way. It is certainly unreasonable to require a complete and precise alternate specification. Nevertheless, some approaches to making possible the evaluation of natural language system capabilities were suggested. Allowable syntactic constructions can be listed, of course. It is virtually meaningless, however, to simply list such constructions as quantifiers, comparatives, and coordinate structures because no one handles them in all of their complexity. Copious examples of essentially different sentences that are allowed should be presented. In addition, gaps and inadequacies of grammars, also illustrated by examples, should be given whenever possible. What a grammar excludes or treats in a less than satisfactory manner is just as important information for evaluation as what it handles nicely. Kuno's description of his predictive analysis grammar (12) is to be cited as an example to be emulated in this regard.

There is a natural tendency, upon observing that two different constructions are provided for in isolation, to assume that the two can be combined in a single sentence, perhaps in a number of different ways. Unfortunately, this is an assumption that is often not well-founded, and originators of natural language processing systems should take pains to address themselves to this point when assessing the capabilities of their systems.

Alternate characterization of a natural language subset can at best aspire to give an interested user or evaluator a rough feel for that language. Questions as to the acceptability of specific sentences can, of course, be resolved only by recourse to the primary mechanism for language specification, the grammar or computer program. For this reason the system designer is under an obligation to make that mechanism available to interested investigators. Detailed and well-organized documentation is also essential if the functioning of a language processing system is to be understood. It is, of course, at best a difficult task to determine by hand how a particular sentence is handled by a given grammar or program. It is much easier to use a working computer program which incorporates an underlying grammar. A directed generation program can provide much help, and an associated parsing program is an even more convenient aid, although in general less efficient with respect to computer time. Such programs should be made available for extensive exploration of a system's capabilities, and, within reason, questions as to the acceptability of specific sentences should be answered by a system designer through recourse to his own computer if necessary.

A final means of natural language system evaluation is possible when that system has both a syntactic and a semantic component, i.e., when a complete system exists for determining the meaning of a given sentence and taking appropriate action. In this case operational tests can be carried out to determine whether users of a language system can effectively use that language to solve given tasks, e.g., question-answering or natural language programming tasks. There are, of course, many variables to be considered in conducting such tests: educational background and system-specific training of the user, difficulty of the assigned problem, quality and difficulty of the solution, for example. Nevertheless, without such attempts at system evaluation it will be impossible to determine whether a given system represents an improvement over its predecessors. Evaluation attempts would be less essential if the natural language systems were all used by significant numbers of people over reasonably long periods of time. Unfortunately, this is not the case. As far as is known, only REL (4), The Lunar Science Natural Language Information System (27), Proto-RELADES (15), and perhaps CONVERSE (11) have attempted user utilization of a question-answering system, and attempts at system evaluation have been meager.

Status of Current Grammars and Systems

The second session was devoted to a survey of individual efforts aimed at producing "computer-oriented" grammars. These were discussed in groups according to the taxonomy previously mentioned. A large number of computer-oriented grammar developers were in attendance and hence able to summarize the current state of their own projects. Other efforts were also noted and commented upon by the person or persons most familiar with them.

One ground rule suggested in session two was that only those grammars be discussed for which some computer implementation of sentence generation or parsing has been written and is still operational. This rule was not always adhered to, because the time at which a program "expires" is usually somewhat vague as is the time at which a large program under development becomes operational. However, for the most part, emphasis was placed on current efforts rather than upon projects which are no longer actively pursued.

Also, no attempt will be made in this report to exhaustively survey all of the recent efforts in the development of computer-oriented grammars, although such an undertaking was the intent of session two.

Because of the limitation of available time, this report cites only the efforts which came up for discussion and assessment of their present status.

Beginning the discussion with context-free and equivalent grammars, it was observed that work on the Harvard Predictive Analyzer program (12) has been terminated and that the system is no longer maintained. The grammar is, of course, preserved in written form, and it lives on in a PL/I version written by Damerau (3). In sheer coverage of English and in the quality of its documentation it has few if any equals, but the inadequacy of structural descriptions it assigns to sentences limits its utility.

Kay's work on parsing systems for rewriting systems was next reported upon because it specializes in context-free and context-sensitive versions in addition to providing parsing for certain general rewriting system grammars. The current parser is called the "Chart Parser" and it was described as a second generation extension of earlier work (10). Current descriptions of the Chart Parser and specific grammars written for it are not yet available, but they are being utilized in the MIND System (8), and documentation in conjunction with that project is in preparation.

Kaplan described a general facility based on the Chart Parser which he calls PARSER or SYNTAX MACHINE and which can be specialized to accomplish a variety of functions. It was claimed that these included not only rewriting system parsing but also augmented transition network parsing and transformational generation and parsing. A paper on this work is to be included in the proceedings of the NYU Symposium on Natural Language Processing (9).

Very little was said about the CONVERSE System (11) in the absence of anyone who has worked on it. Kay reported that his Chart Parser has replaced earlier syntactic components utilized by that system.

REL (4) was described by Dostert as a total software system, a subset of English being only one of the user languages available under it. Others include algebraic, music, and film languages. The entire system (syntax, semantics, data base representation and accessing, etc.) is currently being transferred from an IBM Model 50 assembly language implementation to a GE 135 assembly language version.

A good-sized grammar is being written at the Linguistics Research Center of the University of Texas in conjunction with a German to English translation project (21). The syntactic component used in this work consists of a sequence of context-free grammars, each augmented by agreement considerations with respect to features associated with the nodes. Translation from the structure assigned by one of these grammars to a string that constitutes the input to be parsed with respect to the next grammar is accomplished by a translation mechanism similar to that described by Lewis and Stearns (13).

Winograd reported on the current status of his PROGRAMMAR system for the syntactic and semantic specification of a language (24). The program was reported to be written in a particular version of LISP and work is underway on producing a clearer and cleaner version. Some versions have already been run on CDC and IBM LISP systems. Although the system's syntactic and semantic processing are intertwined, the parsing process is somewhat independent of semantic processing; the first syntactic analysis which meets semantic criteria is taken, and statement of backup paths to be followed must be explicit. This apparently complicates the task of extending a PROGRAMMAR-specified language.

Because of its deterministic formulation, Winograd's syntactic component differs from that of Woods' augmented transition networks. The latter provide a nondeterministic continuation mechanism. Also, they are grammar-specified rather than procedure-specified, as is the case for PROGRAMMAR. In spite of these differences, Woods' and Winograd's syntactic components are very closely related.

Augmented transition networks provide a clean line between syntax and semantics. A single syntactic component has been used by Woods in conjunction with two different data bases (27). Semantic interpretation with respect to those data bases is carried out by the same procedure which was used in the Airline Guide application (25). The syntactic component is, at the moment, better developed than the semantic component. Not all syntactic structures are currently translated. The data bases in question are designed to support defense document retrieval requests and moon rock data query. Results of the former application show that 75% to 80% of the sentences in a corpus contributed by users familiar with the application were parsed after necessary lexical doctoring. Of the sentences accepted, the right parsing was described as often not found, but a "reasonable" parsing was reported to be found 50% to 60% of the time. Implementation of Woods' question answering systems has utilized the Bolt, Beranek and Newman (BBN) LISP System.

Robert Simmons reported the use of a system which he described as "a variant of Woods'", hence presumably also formulated as an augmented transition network. Deep structures are similar to those of Fillmore's case grammars. A second grammar is used to map those deep structures to paraphrases of the original sentences. A CDC LISP System is used as the program's means of implementation.

The NYU String Program (18) is consistent with Joshi's string-based transformational grammars. The program is based on two sequential steps---string decomposition and transformational decomposition of analyzed strings. Work is still in progress on the latter step. The existing program and underlying grammar have been used in analyzing sentences taken from English scientific articles. It was reported that a "correct" parse was included among the first three obtained about 65% of the time. An option is provided for collecting into a single structural description distinct structures which differ only with respect to the assignment of a modifier to a particular host. Resolution of these ambiguities requires further development of selectional restrictions.

Joshi referred to work on imposing conditions on a transformational grammar that insure generation of a recursive language and also to work on relating string-based and tree-based transformational grammars (7). An algorithm for syntactic analysis with respect to the former class of grammars has been developed, and a particular large-scale grammar of English belonging to this class is being written by Danuta Hiz. Only a small part of the grammar has been tested so far, and documentation of the grammar is not expected to be completed until May 1973. It has not yet been decided whether to make use of the NYU String Program in achieving a capability for the syntactic analysis of sentences with respect to this grammar or, instead, to produce a completely independent implementation.

Work on a tree-based class of transformational grammars is being conducted at the IBM Research Center. The class of grammars in question is an extension of the class previously allowed by Petrick (16), and provides for: binary features to be associated with terminal and non-terminal nodes, feature-sensitive and feature-manipulative transformational rules, a general labelled bracket condition of transformational applicability, the expression of supplementary conditions of transformational applicability, and an extended mechanism for specifying desired structural changes. An IBM 360 LISP System implementation is in operation, and a grammar is under development for a question answering system application (17). This grammar currently has nine base component context-free (CF) phrase structure rules, 134 surface structure CF rules, 170 derived constituent structure CF rules, and 37 rather complex transformations. An early version of the grammar was tested using Friedman's transformational grammar tester program (6).

Friedman's program has, in fact been used in developing and testing a number of transformational grammars. There are a number of reasons for this: the program is written in FORTRAN and is relatively exportable, it is well documented, and it was given a degree of generality in both transformational expressability and cyclic control which permits specialization to different theoretical models.

Two past efforts carried out at the IBM Cambridge Scientific Center were discussed by Moyne. These are the Proto-RELADES System (15) and the CUE System (14). The former, designed for use in a library catalog question answering system, utilizes a syntactic analysis specified grammar, i.e., a surface structure context-free grammar and a set of transformations to map the surface structures thus specified into corresponding deep structures. Hence, this linguistic model is seen to be rather different from the usual generative transformational model. The CUE System features more extensive syntactic facilities but no semantic interpretation with respect to a data base. In this syntactic model context-free parsing and analysis-specified transformation are interspersed rather than separated as in the Proto-RELADES System.

Three distinct components are used sequentially in Grenoble University's parser, which is part of a Russian to French automatic translation program (23). The first component is devoted to morphological analysis. It uses a finite state transducer to map strings to strings. The second component uses a CF grammar augmented with context-sensitive features, and a bottom-up parsing algorithm to map the output strings of the first component into surface structure trees. The third component utilizes a tree to tree transducer having the power of a Turing machine to convert surface structures into "tree-formulas" which represent the elementary statements of a sentence and their logical connections. These, in turn, provide the basis for translation to the target language. The system has been tested on a corpus of about 400,000 words, divided into texts of about 3,000 words each.

The University of Montreal incorporates another type of parsing system in its English to French translation program. The Q-System (2), a rewriting system related to that of the MIND System, is used to separately provide models for morphological and syntactic analysis. These are sequentially applied in obtaining a structure which is the basis for English to French translation. Testing of the system on random texts taken from colloquial conversations and from the field of economics is currently in progress.

A number of other efforts related to computer-oriented grammars were either discussed very briefly or merely mentioned. These are represented by appropriate references. No attempt has been made to include in the references all relevant papers; instead, only the most recent papers describing a particular line of research are listed and the interested reader can consult them for references to past work.

Parsing

The third session focused on parsing. Although an attempt was made to elicit information on computational time requirements, the incommensurability of the systems in question all but precluded meaningful comparison. Even more important than the characteristics of the different computers used are the differences in the syntactic coverage of languages employed as well as the differences with respect to the adequacy of structural descriptions assigned. A time figure is not really meaningful without pinning down the sentence, grammar, computer, and parsing algorithm. Times cited for a fifteen to twenty word sentence ranged from ten to as much as 130 seconds.

A few other topics related to parsing were also brought up. The first relates to the question of the relative efficiency of interspersing syntactic analysis and semantic interpretation as compared to sequentially performing those functions. Clearly, it can be no less efficient to allow (but not require) interspersion of these two components. It is not clear, however, just how much semantic interpretation prior to completion of parsing is optimal. This depends, of course, upon the particular syntactic analysis and semantic interpretation components in question. It is interesting to note, however, that for two systems reporting experimentation designed to determine the optimal amount of semantic interpretation to be carried out prior to finishing syntactic analysis, the answer turned out to be none. That is, it was found to be more efficient to delay semantic interpretation until completion of syntactic analysis. Evidently, the time lost in semantically interpreting structures which would later have been precluded for syntactic reasons was greater than the time gained by blocking semantically anomalous structures that would only later have been blocked for syntactic reasons. The two projects in question were the REL effort (4) and the BBN Lunar grammar (27). Results for the former are not so surprising because of the relatively high cost of semantic interpretation due to the use of secondary storage for data base representation; however, the BBN Lunar grammar used a single level of storage.

Although a few new topics related to parsing were identified, very little in the way of results was reported. These included the following: (1) "errorful" parsing, i.e., parsing strings generated by a given grammar plus an additional noise component; (2) parsing algorithms that proceed from given anchor points in a sentence rather than proceeding strictly left-to-right or right-to-left; (3) parsing with respect to phonological component grammars, and (4) natural language subset extension, i.e., incremental modification of a grammar both directly and through statements in the language of that grammar.

To describe the current situation with respect to developing large-scale grammars and syntactic analyzers, one can say that this activity is certainly not currently being undertaken by many linguists, whose typical practice consists of examining isolated but hopefully key linguistic phenomena and proposing modifications to linguistic theory on the basis of that examination. Putting together a large set of rules (often, any explicit rules) is very unfashionable, as is the detailed and complete specification of a proposed class of grammars (i.e., a linguistic theory). Nevertheless, there is some ongoing research in these areas, and one could hazard the identification of three distinct types of models being used. The first is some kind of a phrase structure grammar, often a context-free grammar frequently augmented by binary features and even by a set of "transformations" which map surface structure into a "deeper" representation. This first model is not based on an underlying transformational grammar.

A second class of models consists of generative transformational grammars. There have been more rule testing systems than parsing systems developed for these grammars. The parsing systems that have been developed formally resemble to some extent the "transformationally" augmented variant of the first class of models; they are, however, usually slower because of the necessity that certain nondeterministic continuations be followed.

The third class of grammars includes such efforts as those of Woods (26), Winograd (24), Bobrow and Fraser (1), and Thorne, Bratley, and Dewar (22): augmented transition networks and their related efforts. Proponents of this class of models hold forth the hope that their devices will prove useful not only as a primary means of linguistic description but also as a useful formalism for specifying parsers of the other two classes of grammars. They have, indeed, been used to construct a parser for a given context-free grammar; however, more general results, in particular the construction of a parser for a given transformational grammar, have not been obtained.

One observation on these three classes of grammars concerns the types of structures they assign to sentences. They all are capable of assigning tree structure (with associated features) to sentences, and, as used by their proponents, they all refer to deep and to surface representations. Not long ago the view was commonly held that although there were considerable differences between the models in use in computational linguistics, the differences between the structures they assigned to the same sentences were not so great. This was, in any case, the consensus expressed at the Ames Seminar on Text Processing Research of January 1970. It seems less true today, however. We have differences between structures assigned by case grammars and non-case grammars, by generative semantic and interpretive grammars, etc. Indeed, the terms deep and surface structure are relative rather than absolute, and the term deep structure in particular may refer to very different types of structures in different linguistic theories.

A final question about the grammars identified in this report concerns the degree of their capabilities for specifying subsets of a natural language and assigning acceptable structures to the sentences specified. This question was raised during both sessions, and, although it could not be answered clearly, a number of suggestions were offered to grammar writers and parsing system implementors. If followed, these suggestions would probably make an evaluation of those capabilities possible.

References

1. Bobrow, D.G., and B. Fraser. "An Augmented State Transition Network Analysis Procedure." In: Proceedings of the International Joint Conference on Artificial Intelligence, 557-567. Washington, D.C., 1969.

2. Colmerauer, A., J. Dansereau, B. Harris, R. Ketteredge, and M. Van Caneghem. An English-French MT Prototype. Montreal: Projet de Traduction Automatique, University of Montreal. (Text of a paper presented to the 1971 International Meeting on Computational Linguistics, Debrecen, Hungary, September 1971.)

3. Damerau, F. J. "Automatic Parsing for Content Analysis." Communications of the ACM 13(1970):356-360.

4. Dostert, B.H., and F.B. Thompson. The Syntax of REL English. (REL Report No. 1.) Pasadena, Calif.: California Institute of Technology, September 1971.

5. Fabry, R.S. "A COMIT Generating and Parsing Program for Left-to-Right Subscript Grammars with Discontinuous Constituents." In: W.A. Kalenich, ed. Information Processing 1965. Proceedings of the IFIP Congress 65. Vol. 2, 529-530. Washington, D.C.:Spartan Books, 1966.

6. Friedman, J., with T.H. Bredt, et al. A Computer Model of Transformational Grammar. New York: American Elsevier, 1971.

7. Joshi, A.K. "A Brief Survey of Some Mathematical Results Relevant to Natural Language Processing." In: R. Rustin, ed. Natural Language Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1972 (forthcoming).

8. Kaplan, R. The Mind System: A Grammar-Rule Language.(RM-6265/1-PR.) Santa Monica, Calif.: The Rand Corporation, April 1970.

9. Kaplan, R. "A General Syntactic Processor." In: R. Rustin, ed. Natural Language Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1972 (forthcoming).

10. Kay, M. "Experiments with a Powerful Parser." In: Proceedings of the Second International Conference on Computational Linguistics. Grenoble, August 1967.

11. Kellogg, C., J. Bueger, T. Diller, and K. Fogt. "The Converse Natural Language Data Management System: Current Status and Plans." In: J. Minker, and S. Rosenfeld, eds. Proceedings of the Symposium on Information Storage and Retrieval, 33-46. College Park, Md.: University of Maryland, April 1971.

12. Kuno, S. "The Current Grammar for the Multiple-Path English Analyzer." In: Mathematical Linguistics and Automatic Translation. Report No. NSF-8. Cambridge, Mass.: Computation Laboratory, Harvard University, 1963.

13. Lewis, P.M., and R.E. Stearns. "Syntax Directed Transduction." Journal of the ACM 15(July 1968).

14. Loveman, D.G., J.A. Moyne, and R.G. Tobey. "CUE: A Preprocessor System for Restricted Natural English." In: J. Minker, and S. Rosenfeld, eds. Proceedings of the Symposium on Information Storage and Retrieval, 47-59. College Park, Md.: University of Maryland, April 1971.

15. Moyne, J. Proto RELADES: A Restrictive Natural Language System. (TR-BPC3.) Cambridge, Mass.: IBM, October 1967.

16. Petrick, S.R. A Recognition Procedure for Transformational Grammars. Unpublished Ph.D. dissertation, M.I.T., 1965.

17. Petrick, S.R. "Transformational Analysis." In: R. Rustin, ed. Natural Language Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1972 (forthcoming).

.8. Sager, N. "The String Parser for Scientific Literature." In: R. Rustin, ed. Natural Language Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1972 (forthcoming).

9. Schank, R.C., and L.G. Tesler. "A Conceptual Parser for Natural Language." In: Proceedings of the International Joint Conference on Artificial Intelligence, 569-578. Washington, D.C., 1969.

0. Simmons, R.F. "Natural Language Question-Answering Systems: 1969." Communications of the ACM 13(1970):15-30.

1. Stachowitz, R. "Requirements for Machine Translation: Problems, Solutions, Prospects." In: Winifred P. Lehmann, and Rolf Stachowitz. Feasibility Study on Fully Automatic High Quality Translation. Final Report, 1 February 1970--30 June 1971 (Report No. RADC-TR-71-295.), 59-96. Griffiss Air Force Base, N.Y.: Rome Air Development Center (IRDT), December 1971.

2. Thorne, J.P., P. Bratley, and B. Dewar. "The Syntactic Analysis of English by Machine." In: D. Michie, ed. Machine Intelligence 3, 281-309. New York: American Elsevier, 1968.

3. Vauquois, B. "Intermediate Language for Automatic Translation." In: R. Rustin, ed. Natural Language Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1972 (forthcoming).

24. Winograd, T. Procedures as a Representation for Understanding Natural Language. (MAC TR-84.) Cambridge, Mass.: Project MAC, Massachusetts Institute of Technology, February 1971. [M.I.T. Ph.D. dissertation, 1971.]

25. Woods, W. "Procedural Semantics for a Question-Answering Machine." AFIPS Conference Proceedings. Vol. 33, Part 1. 1968 Fall Joint Computer Conference, 457-471. Washington, D.C.: Thompson Book Co., 1968.

26. Woods, W. "Transition Network Grammars for Natural Language Analysis." Communications of the ACM 13(1970):591-606.

27. Woods, W., and R. Kaplan. The Lunar Sciences Natural Language Information System. (BBN Report N. 2265.) Cambridge, Mass.: Bolt, Beranek and Newman, Inc., September 1971.

MACHINES AND SPEECH

Franklin S. Cooper, Chairman

Introduction

Speech, as a topic for the Conference on Research Trends in Computational Linguistics, was included in the agenda because of its emerging relevance to the field. This factor set the tone for the opening discussions within the group and was largely responsible for the fact that most of the group's attention was spent on the two topics for which the overlap with the subject matter of conventional computational linguistics was most clearly evident. There was tacit agreement that many of the problems that concern speech researchers and for which they regularly use computers would not be of interest to most of the conferees. Thus, there was little discussion of such current speech areas as the physiology of speech production and audition, perception of speech and speech-like stimuli, cross-linguistic studies of speech and universal phonological theory, children's language acquisition, speech and hearing disabilities, etc.

Instead, the discussion centered on speech production by computers in the context of reading machines for the blind and on speech recognition by computers as a central problem in designing speech understanding systems. There was in addition some discussion of the research tools needed both for research in depth at major research centers and for graduate level training in a larger number of university laboratories. The Chairman, in his report to the plenary session, dealt only briefly with the state of the art in speech research but put primary emphasis on research opportunities, covering some areas in addition to those which had received attention in the group discussions.

Speech as a Part of Computational Linguistics

Computational linguistics, as defined by past usage, has dealt mostly with written rather than spoken language. This is mainly an historical accident; nevertheless, the time has come to examine the areas in which speech may now be considered a part of this field. There are, indeed, new factors to be considered: one is that the domain of automated language processing is, in practice, being extended to include systems that generate speech as an output and that accept it as an input. Another reason, based in theory, is a growing awareness of parallels between low level processes of speech production and perception and higher level processes involving syntax and semantics. This has come about mainly as a consequence of psycholinguistic experimentation on human language processing at all levels, although most of the effort---and the more penetrating methodologies---have been applied at the level of speech. Thus, research on human processing of language, speech included, can serve to suggest areas and methods that computational linguists may wish to explore.

This is not to say that all of speech research ought to be coopted into computational linguistics: much that is merely descriptive or that is concerned with the technology of voice communication has little to offer or to gain in return. A reasonable criterion might be that there should exist a mutual interdependence between the speech processes and the higher level processes involved in the same overall automated language operation. This, at least, was the basis we used in choosing topics for discussion, and in considering practical applications. The nature of the interdependence will become evident in the following discussions of reading machines and of speech understanding systems.

Speech as an Output of Automated Language Processing: Reading Machines for the Blind

Most of the familiar instances of speech from machines---telephone directory assistance, airline announcements, etc.---are essentially uninteresting to computational linguists. They involve such limited vocabularies and fully defined syntactic structures that adequate speech can be had simply by using prerecorded words and phrases or their synthetic equivalents. The more general problem of generating speech from unrestricted text is only now approaching solution, primarily in connection with reading services for the blind. Indeed, many of the problems that face any automated speech output device of a non-trivial design can be described and analyzed by detailed consideration of this single application. The practical problem of providing such a service has many additional aspects, nonlinguistic as well as linguistic. Here we shall undertake no more than a sketch of component operations and problem areas.

The primary function of a reading machine for the blind* (so called because of the many attempts to build simple, portable devices) is to provide blind people with adequate access to the full range of printed materials that sighted people read. An approach that has been tried repeatedly during the past sixty years is to use a photoelectric probe that converts letter shapes into sound shapes which the blind user is expected to learn. This task, it is argued, is no worse than learning a foreign language; indeed, it should be easier, since the lexicon and syntax are those of English. In practice, learning is laborious and reading rates are disappointingly low---comparable to Morse Code and roughly an order of magnitude less than listening to spoken language. The reason for the superior performance of spoken language is by now

*Two recent reviews that deal with the general problems of sensory aids for the blind (including reading machines) and that give extensive references to the literature are "Electronic Aids for the Severely Visually Handicapped" by Jonathan Allen in Critical Reviews in Bioengineering, D. Fleming (Ed.), Chemical Rubber Co., 1971, pp. 139-167, and "Sensory Aids for the Blind: A Challenging Problem with Lessons for the Future" by P.W. Nye and J.C. Bliss, Proc. IEEE, 58, 1878-1898 (1970).

quite clear: speech is a highly condensed encoding of the message, whereas the letter sounds provide only a serial encipherment into acoustic form and so are far less efficient carriers of linguistic information.

Both theory and experience suggest, therefore, that high performance by a reading machine requires that it be able to speak plain English---or whatever language is being used. The complexity and size of such a machine are such that it will have to be, for the present at least, a central facility that records tapes on request, or possibly serves a remote user who is linked to it by telephone. Many of the practical problems are therefore those inherent in the organization, financing, and administration of any sizeable service function such, for example, as a time-sharing computer system.

In the operation of a reading service center, the first step is to obtain a machine-readable alphanumeric tape from the printed text that was requested by the blind user. Sometimes compositor's tapes (from which the book was printed) will be available, but usually the printed page must be converted by optical character recognition (OCR) machines, or by manual keyboarding. [This problem of entering data from the printed page is shared with many projects in computational linguistics. At the present time, there are no suitable OCR machines available as standard equipment, and only a few companies are prepared to supply special machines capable of reading proportionally-spaced characters in multiple fonts. Service bureau facilities for reading conventional printed text are likewise limited and expensive. The reason for this situation lies only partly in the technical difficulties; mainly, there is less customer demand for such devices than for simpler and cheaper machines designed to read at high rates the specialized, uniformly-spaced characters from credit cards and business documents. It may well be that the needs of computational linguists, as well as those concerned with reading centers for the blind, can only be met through a development project aimed at their special needs, i.e., for moderate accuracy, moderate speed, and reasonable cost in a machine with enough virtuosity to recognize the commonly used fonts and to scan bound volumes. Whether or not such a project falls within the scope of computational linguistics remains open to question. It could nevertheless be so useful for text input to computers that a good case can be made for it.]

The next step, once the text is available to the computer in alphanumeric code, is to arrive at the pronunciation of each word in terms of some appropriate phonetic notation. Due to the nature of English orthography, no simple set of rules can be employed to derive an acceptable guess at the proper transcription for all English words, though there are spelling regularities that may, perhaps, be exploited to advantage. Hence, some kind of pronouncing dictionary (in machine-readable form) is essential. Two general approaches are being tried: (1) There is the straight-forward, pragmatic one of storing the phonetic equivalents for every word of a large lexicon, including separate entries for most root-plus-affix combinations. This allows the specification of inherent lexical stress, as well as of a code indicating the usual syntactic role of the word. A dictionary of this kind with approximately 150,000 entries, about equivalent to a desk-type dictionary, is easily accommodated on a single IBM 1316 disk pack. (2) With a more analytic approach, considerable savings in dictionary size, perhaps ten-to-one for very large dictionaries, may be achieved by attempting to break the input words into their constituent morphs. However, word pronunciation is not a simple function of the pronunciation of its constituent parts: for example, a suffix may shift the placement of primary stress and force a change in vowel quality. In any case, the dictionary must contain phonetic equivalents for the full set of morphs and for a substantial number of frequent words that violate spelling-to-pronunciation rules. Research problems of considerable interest are involved in such an approach: for example, the development of a set of pronunciation rules based on an underlying representation of English morphemes, working out algorithms for word decomposition, finding rules for the placement of lexical stress in words that are reconstituted from their morphs, and inferring syntactic roles that such words can play.

Once a canonical phonetic representation has been obtained for each word in the text, it is necessary to employ a set of phonological rules which determine how the sequence of words which constitute an integrated sentence should be spoken. Perceptual experiments with concatenated words, using recordings of single words spoken in isolation, have demonstrated the importance and the extent of the sentence-level recoding of spoken word strings. As an example at the segmental level, a rule transforms the phonetic sequence [d, word boundary, y] into [ǰ], which means that the normal segmental realization of the sentence "Did you go?" is actually [dIǰugo]. While effects of this sort are familiar, the exact form which the rules take is not known and the inventory of such segmental rules is far from complete. Finding these rules is a challenging research problem and one to which computational methods can make an important contribution, especially in testing the reliability and range of application of proposed rules.

Of more importance than such segmental rules are phonological rules that determine the temporal organization of the acoustic output and the fundamental frequency of vocal cord vibration as a function of time. Segmental durations and intonation contours are influenced by a number of factors, especially the syntactic structure and stress pattern of the sentence. Indeed, some sort of syntactic analysis is essential for the synthesis of a satisfactory, spoken sentence. In some degree, structure can be inferred from orthographic punctuation, and rules based on punctuation are sometimes sufficient. Nevertheless, better methods are needed. This dependence of speech quality on syntactic structure is perhaps the major area of overlap between the reading machine problem and conventional computational linguistics, and a promising target for further research.

At this point in the process of generating speech from written text, the computer has assembled a phonetic string that has had appropriate allophonic and stress adjustments and that has been marked for intonation and juncture. It remains to convert this phonetic description into control signals that will operate a hardware synthesizer, or its simulation in software. Two general methods of speech synthesis by rule are currently utilized: (1) For a terminal-analog synthesizer, the rules manipulate acoustic variables directly, e.g., sound source type and pitch, formant frequencies, and intensities. In this case, the rules for synthesis begin by consulting tables for the canonical form of each phone, then computing the necessary contextual adjustments (corresponding approximately to coarticulation in human speech). Typically, there are about a dozen parameters that are used to control a hardware synthesizer; these are specified at regular intervals of about ten milliseconds, requiring a total output bit rate of about 4800 per second. (2) For an articulatory-analog synthesizer, the rules typically manipulate articulatory variables such as the positions and shapes of simplified models of tongue, lips, velum, larynx, etc. The resulting shapes (of the model vocal tract) are then used to compute an acoustic output, or the control signals for terminal-analog hardware. (It should be pointed out that substantially more work has been done on terminal-analog models and rules which currently provide a more intelligible output than do articulatory implementations; however, articulatory models are improving rapidly and are of greater theoretical interest because rules of coarticulation may ultimately be built in as automatic constraints. In addition, modelling of this kind may lead to a better understanding of the physiology of speech production; indeed, this is an area of promise for future research to which the present discussion will return.

The intelligibility of the synthetic speech currently produced by rule with terminal-analog systems is surprisingly good. Tests of consonant-vowel words and nonsense syllables synthesized by rule have shown that listeners are able to identify them correctly 95% of the time. Systematic testing for longer words and sentences has not been done, but interviews and informal tests have been carried out with blind students who have listened to chapter-length recordings of textbook materials. Few words are missed and overall comprehension is high. Nevertheless, it is clear that the treatment of consonant clusters and of stress and phrasing needs improvement and often seems unnatural. Future improvements in the rules will depend on careful analyses of natural speech and the systematic manipulation of synthetic speech to test the perceptual relevance of proposed changes in the rules.

The evaluation of synthetic speech in terms of its real-life utility to blind individuals is a major task. The speech itself needs to be considered along two dimensions---intelligibility and naturalness. The voice quality of current synthesizers is not as natural as would be desirable, despite efforts to improve it. This may imply that we are ignoring some critical variables such as glottal source irregularities or, alternatively, that naturalness will not be achieved until the synthesis rules are improved sufficiently to avoid all of the conflicts between acoustic cues and message content. It is interesting to note ---and ultimately encouraging---that it is the rules in their present form and not the synthesis hardware that is at fault; in fact, extremely good synthetic speech---indistinguishable from the spoken version---can be made by meticulous hand adjustment of the control parameters.

But complete naturalness may be too much to expect for speech that is synthesized by fully automatic methods. Evaluation must then deal with the question of how useful the product is for, say, the blind student in preparing his lesson assignments. Its principal advantage over natural speech is that he should be able to get the material he wants when he needs it, not some weeks or months later as often happens with recordings by volunteers. The computer can read tirelessly and faster than a human, though its actual performance will depend on how well the service function is organized. Good intelligibility of the synthetic speech is, of course, essential but this may not be an adequate criterion. It might be that listening to synthetic speech imposes so much perceptual load that comprehending and remembering the content would be excessively difficult; hence, comprehension tests and measures of fatigue are more likely to be relevant than intelligibility tests in evaluating the practical usefulness of computer-generated speech. Evaluations of this kind are being started, in parallel with efforts to make the synthetic speech sound more natural.

This sketch of the reading machine problem has pointed to some of the areas of interdependence between speech research and computational linguistics. Thus, many problems of dictionary management are shared.

Spelling-to-sound rules and algorithms for decomposing words into constituent morphs would reduce the size of the dictionary needed for a reading machine, just as comparable algorithms for syllabification do for automated typesetting. Likewise, reliable methods for reconstituting words and for inferring their usage would be useful to either a machine that must read them aloud or to one that is composing written responses. But the common ground is most evident at the level of syntactic analysis. An efficient general purpose parser is almost equally necessary for properly rendering a sentence into spoken form, or for inferring its content from its written form. For the present, reading machines must depend on explicit punctuation and a pseudo parse of some kind; perhaps, short-cut methods that yield good speech would also have practical application to the automatic punctuation of synthetic written responses. The existence of this common ground implies research opportunities on the interrelations of spoken and written language---another topic for later discussion. But before dealing with areas for research and practical application, the report will give an account of the discussions on speech understanding systems.

Speech as an Input to Automated Language Processing: Speech Understanding Systems

Most of the discussions in this session centered on Advanced Research Projects Agency (ARPA)-sponsored projects on speech understanding systems, underway for about six months. Five major research groups, already engaged in other ARPA-supported work, are involved. As an initial step, a study group assembled from these projects analyzed the problems and prepared a set of objectives, specifications, and plans.*

The overall objective is to develop one or more demonstration systems, primarily to show that the technology now exists---though it

*"Speech Understanding Systems: Final Report of a Study Group". Published by Computer Science Department, Carnegie-Mellon University, Pittsburgh, Pa., May 1971.

is scattered---to make such an undertaking feasible. Each of the major research groups is undertaking a task of its own choosing, but with the expectation that a cooperative effort on some one of these projects, or an entirely different one, will emerge. Definition of the tasks was considered to be a crucial point in setting the level of difficulty and the chances for success. Indeed, the long background of failures to solve the "speech recognition problem" was in part responsible for the ARPA decision to undertake a related, but more manageable, program.

The machine recognition of speech has been a persistent challenge for at least the past twenty years. Several ways have been found to recognize spoken digits, or even a few dozen words when they are spoken singly and carefully. But the general problem, usually put in terms of a phonetic typewriter or a system for automatic dictation, has remained elusive. The much larger vocabulary that is required, the necessity of dealing with connected speech, and the need to accommodate a number of different speakers have all posed severe difficulties. Moreover, as more has been learned about the nature of the speech signal and its understanding by humans, the clearer has become the magnitude and complexity of the recognition problem. It should be noted that most attempts thus far to deal with the general problem have used a "bottom-up" approach, i.e., one in which phonetic elements, words, and sentences, are all found by successive analyses based on the acoustic signal. The speech understanding systems projected by the ARPA Study Group differ in two important ways: constrained objectives make the problem more manageable, and reliance is not placed on bottom-up analyses.

Speech understanding, in the present context, means essentially that a computer, when told to do something, will "understand" well enough to take the correct action or will ask for clarification. If, for example, the computer has control of a robot, then a command to put the red block in the box can lead to an easily confirmed performance, provided there is a red block, and a box that is larger than the block; otherwise, it should lead to a question or an error message. Or again, if the computer contains a file of information about moon rocks, it should be able to answer inquiries about the numbers, sizes and chemical compositions of those rocks, whether the query is phrased as a question or as a demand for data. Obviously, tasks of this kind are multidimensional and can be constrained in ways that will, in fact, determine their difficulty. The attempt has been to choose tasks---more or less like the two mentioned---that are constrained in such a way as to make them manageable, but not to the point of making them trivial. A practical payoff was not considered a mandatory requirement.

Performance in the task situation not only limits the number of possible responses to the voice input, it also provides additional bases for analyzing it. Thus, limitations on the vocabulary, on the syntax that is permitted for the task, on allowable operations to be performed, and predictions of the probable behavior of the person speaking---all provide bases for making hypotheses about the actual message carried by the incoming acoustic signal. Indeed, it is on such lexical, syntactic, semantic, and pragmatic "support" that the research groups are putting their hopes and focusing their efforts; although the acoustic signal is not being ignored, it is receiving far less attention than it has in past efforts to solve the general recognition problem---quite possibly, less attention than it will need to receive if effective use is to be made of "top down" or "side in" approaches.

The ARPA Study Group's final report discusses at length the kind of support, and the nature of the problems, to be expected at the various levels. Interdependence across levels is a characteristic of the entire undertaking, and there is much overlap with the domain of conventional computational linguistics. There are challenging differences, too: neither sentences nor words are well formed, as one would expect them to be in written text; moreover, the need for live interaction between user and computer means that the computations must be done in real time or less, much of it while the sentence is still being spoken.

Only at the very lowest levels, where parameters are being extracted from the acoustic signal and are used to guess the phoneme string, are the problems wholly within the domain of speech research. Here, although engineering problems of pitch derivation and formant tracking are not trivial, the major difficulty lies in inferring the phonetic string---a difficulty that may be inherent in the nature of speech itself. It is clear that in the production of speech there is much restructuring (or encoding) of the segmental units to achieve a compact, smoothly flowing output; that is, there is much overlap of the gestures for units that are themselves sequential. It is, indeed, the rules for this coarticulation that provide a basis for speech synthesis by rule. But the rules we use are generative rules; the inverse rules, to the extent they exist, are largely unknown except for those phones and contexts where coarticulation is minimal and the "code" can be said to be transparent. A general paradigm, proposed by Stevens some years ago and labelled "analysis-by-synthesis", uses heuristics to guess the phonetic string and then confirms by synthesis, or uses error signals to make a better guess. An alternative, referred to by the Study Group as "hypothesize-and-test", looks for transparent places in the code (and for other acoustic landmarks) to generate a much less complete hypothesis about the phonetic string, and then brings to bear information from higher level processes. An example would be to look for a word that begins with strong, high frequency noise, then guess the word to be "six" if it has one syllable, or "seven" if two, provided one has other reasons to expect a spoken digit.

Obviously, there are significant research problems at the speech level, as well as practical difficulties, for a speech understanding system: can a set of rules for analysis be devised? or even a partial set that would have practical utility? What kinds of heuristics will be most helpful if one resorts to analysis-by-synthesis? What relative reliance should be placed on extracting as much information as possible from the acoustic signal versus having to depend on support from higher linguistic levels? Underlying these questions is the assumption that present knowledge of the acoustic cues is reasonably complete; in fact, much remains to be learned about the cues that operate in consonant clusters and in connected speech, both careful and casual.

At the next higher level, the principal problem is how to convert the string of phonetic elements---perhaps incomplete, and certainly error-prone---into a string of words, which may also have intervening gaps. Word boundaries are not at all evident at the acoustic level, and so do not appear in the phonetic string. Segmentation into lexical elements must then proceed mainly by matching to strings that correspond to words in the task vocabulary. Often the low-level analysis will have generated more than one possible candidate for each phone, omitted an element, or supplied a wrong entry. The matching operation must somehow avoid the combinatorial explosion that could readily occur if there were several options for each of a string of phones. Omissions and errors pose obvious additional difficulties. There is, therefore, serious need for support from higher levels as well as efficient analysis at the phonetic level.

The construction of sentences from a partial and "noisy" string of words can draw support from whatever restrictions the task may impose on the syntax, or from information about the location of syntactic boundaries that can be inferred from the suprasegmentals found in the acoustic signal. The latter relationship is essentially the inverse of the dependence of synthesis on syntactic information to assign stress and intonation and thereby generate speech that sounds natural. There are other problems at the sentence level that have their counterparts in parsing written text. They differ, though, in that analysis of the spoken "sentence" will often have to deal with intrusive hesitation sounds, with non-well-formed or incomplete sentences, and with both backwards and forwards parsing from a starting point somewhere within the utterance. There are obvious research problems of considerable importance in devising methods for parsing under these conditions.

Interpretation of the sentence as a basis for action can rely only in part on the output of the syntactic analysis, since that output is likely to be faulty or ambiguous. Good use will need to be made of the semantic constraints imposed by the task, and of pragmatic information about the behavior to be expected of the human operator.

Except for such help, the task of interpretation has all the complexities inherent in the content analysis of written text. If the appropriate response from the computer is an answer to a user's question, then various cognitive functions must also be performed and a sentence must be generated that is appropriate in both content and form. The response itself could be in either written or spoken form, with the latter making use of techniques for text-to-speech conversion developed for reading services for the blind.

It will be evident that speech understanding systems are involved with language at all levels, and pose many problems that should interest computational linguists. The differences between these problems and more familiar ones reflect deep differences between oral and written language. Even so, this account has probably understressed the pervasive influence of speech at all levels of the speech understanding problem, and not merely those levels to which the term is usually applied.

Research Areas With Speech Involvement

The preceding discussions of speech outputs from computers (for reading machines) and speech inputs to computers (in speech understanding systems) have exposed a number of important areas for research in computational linguistics. Some of the basic problems that were mentioned in passing deserve additional consideration.

Translating Between Written and Oral Language

Although it may be overdramatizing the differences between written and oral language to speak of translating the one into the other, it may nevertheless suggest a useful point of view in reexamining some old problems and considering some new ones. We have seen that the research on reading machines for the blind makes explicit some written/oral differences that often pass unnoticed. Ideally, a reading machine should convey all of the useful information that is on the printed page. This may include much more than the bare text we have so far considered, even without taking account of pictures and diagrams, a task far beyond present capabilities. The well printed page uses many typographic devices to organize and modify the literal text: punctuation marks are so commonly used that we think of them as a part of the text, though their realization in sound follows very different rules than those applied to the letter; capitalization, too, is widely used, and for a variety of purposes; then there is the judicious use of different type fonts, sizes, weights, and leadings; and paragraph indentation is only one of the devices used to indicate breaks, subordinations, listings, etc.

All of these carry valuable information to the eye. How, and to what extent, can this information be "translated" for easy use by the ear? Not, one would hope, by overt description of the typography. The commonest forms of punctuation have acoustic reflexes that are fairly simple and regular. Are there comparable acoustic signals for other graphic symbols? The inverse transformations involved in writing are no easier, though they are more familiar: thus, oral questions, statements and exclamations are indicated by their standard typographic signs; likewise, emphatic stress can be signalled by italics. But how does one convey a note of incredulity or petulance without resorting to bald description? We know that the skillful writer can do it; perhaps, then, ways can be found to convey typographic messages of comparable subtlety to the blind listener in a reasonably graceful way. As a practical matter, when one is converting a compositor's tape into speech, something must be done with each of the graphic signals that it contains.

If compositors' tapes seem to carry too much information about graphic details, they can be accused equally of carrying too little of the essential information about the sentence. Where, for example, is sentence stress to be placed during synthesis? Where are the breaks into breath groups for a long phrase? And just how should the intonation be handled? But it is hardly correct to say that the information---or most of it---is not provided; rather, it is buried in the structure, which is why good speech synthesis will be so dependent on adequate parsing. This assumes though that full knowledge of structure is a sufficient condition for making speech that sounds natural. Even if true, much has yet to be done about finding the rules by which sentence structure can be converted into sentence suprasegmentals. Rules for the inverse transformation from suprasegmentals to structure---to the extent that they exist---could be extremely helpful in supplementing the phonetic analysis in a speech understanding system. A related problem that dips more deeply into conventional speech research is the relationship between stress and intonation (as linguistic entities) and the relevant dimensions of the acoustic signal. These relationships are known only in part.

The written sentence is, nevertheless, incomplete in at least some respects: witness sentences that are ambiguous in written form but not in spoken form, because the real-world context is missing in the one case but often is indicated (by stress and intonation) in the other. One can find ambiguity at the lexical level, too---witness homographs and homophones---though the ambiguity in one mode of expression is usually resolved in the other. It might be interesting to explore the conditions for ambiguity wherever it occurs.

These are a few of the translation problems that one would encounter in dealing only with English. It is easy to see, from the different structures and orthographies of other languages that different---but probably not fewer---problems would arise with them. Finally, the general problem of working back and forth between oral and written versions of the same language will be one of increasing concern to computational linguists as automated language processing becomes more and more involved with speech as its input and output modes.

Modelling Speech Production

We have seen that the rules by which speech is synthesized necessarily deal with the higher level processes that are the normal domain of computational linguistics. Although the models of speech production to which those rules apply lie more nearly in the usual realm of speech research, the models must operate from higher level control signals. Hence, the development of speech models and the organization of their control signals is an area of research that is relevant to computational linguistics as well as important in itself. Additional reasons are that the processes of speech production parallel those at higher levels in interesting ways, and that experimental methods for probing the lower level processes are well developed.

The process of human speech production has several distinguishable subprocesses, organized hierarchically into levels. Parallels with the levels of linguistic processing are based mainly on operations that restructure the intended message to make it more compact and to put it into linear form, an obvious requirement for eventual output as a time-ordered acoustic signal. Speech has the additional feature that its processes change mechanisms on the way down; implementation is no longer done at all levels by neuromechanisms, but must include signal transformations in proceeding from nerves to muscles to gross movements and to sound generation. Thus, some of the restructuring, or encoding, that we find in the spoken message is a consequence of interface requirements. Linguists usually stop---perhaps with good reason---when they have specified the phonetic string (or its equivalent in the form of a sequence of feature matrices), leaving actualization of the message to a human speaker.

What remains to be performed---and to be modelled---are the successive conversions that lead eventually to speech: (1) The phonetic string must be grouped into pronounceable units and converted into a pattern of neural commands to muscles. This involves collapsing the string of linguistically discrete elements into an articulatory unit of about syllabic size and the temporal coordination of a substantial number of neural signals. Some of them may be crucial for maintaining phonemic distinctions (and some merely accessory), or all may be directed at achieving a particular target performance. The organization of these unit gestures is a topic of very lively interest in current speech research. (2) The pattern of neural impulses activates the muscles of articulation in a process that may be quite straightforward, or may involve gamma-efferent feedback loops in important ways---another topic of current interest. (3) In either case, the gesture in neural form is converted, subject to muscular and mechanical restraints, into gross movements of the several articulators, and these in turn determine the configuration of the vocal tract and its acoustic excitation. This involves encodings of quite a different kind from those mentioned above, but often quite extensive; they account at least for a mechanical component in coarticulation. (4) Finally, excitation and configuration determine uniquely---but not simply---the acoustic waveform of speech.

Efforts to model these processes have typically worked upstream from the acoustic level, usually dealing with a single conversion. Thus, the work of Fant and others has given us a good grasp of how to convert changing articulatory shapes and excitations into changing acoustic spectra. X-ray movies and sound spectrograms are only two of the experimental methods for exploring and testing these conversions. The relationships between muscle contractions and articulatory movements are under intensive investigation, using electromyography both to measure muscle activity and to infer neural signals, and using x-rays and spectrograms to observe and infer the resulting movements. Efforts are being made to describe the organization of gestures in motor command terms, with verification to be provided by measurements on muscle activity, configurations, and sound.

Computer methods have been used to good effect in both the experimental work and in modelling conversions at the lower levels. They have been used also to good effect, but in quite different ways, in speech synthesis by rule. Thus, the terminal-analog type of synthesis by rule bypasses all the intermediate stages and operates directly from an input phonetic string to the output speech waveform. The articulatory type of synthesis by rule makes a lesser leap from phonetic string to articulatory gestures, then uses level-by-level models to get to the acoustic output. The obvious goal is good modelling of the conversions at <u>each</u> level, confirmed by direct experimental measures wherever that is possible, and also by the synthesis of natural-sounding speech when the models are used in tandem.

Interfacing Speech to Phonology

It might appear from the preceding discussion that the processes of speech begin where those of linguistics usually end, i.e., with the message in the form of a phonetic string or the equivalent sequence of feature matrices. However, the phones and features of the speech researcher are not---or not necessarily---those of the linguist, since they are defined by different operations. It is the resulting mismatch at this level that poses one of the major problems in modelling the total process of generating spoken language.

When linguists use labels such as "compact" and "diffuse" for features, there is no implication that the features will convert directly into components of a pattern of neural commands to muscles; when such labels as "voiced" and "voiceless" are used, the differences in operational definition are concealed, though they are no less real. The differences have their origins in the dissimilar approaches taken by linguists and speech researchers. The latter usually try to work with real mechanisms and models of processes whereas linguists more often concern themselves with relationships that they can formalize as rules, though exceptions are to be found on both sides.

The interface problem, then, has two different complexions: if linguists' rules really reflect underlying mechanisms and processes---a claim that linguists rarely make for them---and if current speech models prove to be tenable then a conversion is surely possible and finding it becomes an important research goal. But it is conceivable that the linguists' rules are wide of the mark as to processes, however useful they may be as descriptive devices. It would not then be possible to find the "real" conversion, though the search for it might make clear the directions in which phonological theory ought to move. In any case, the problem is inescapable in some guise when automated language processing must operate across the boundary between speech and phonology.

Converting Generative Rules Into Analytic Rules

The discussion of modelling speech production, including the special case of interfacing speech to phonology, has all been generative and most of the models that are concrete and believable are likewise models of production processes. This does not imply that perception has been less studied than production, but only that the research has yielded a more coherent set of relationships for the latter.

There may be a good theoretical reason why this is so: the production process includes important operations that are in principle irreversible---irreversible in the same sense that a drainage system would be irreversible, i.e., water does not run uphill and, if it did, it would not know which way to go at the confluence of two "downhill" streams. To the extent that speech perception is organized in motor terms and shares these irreversible operations, it cannot be expected to provide a model for straightforward analytic rules. Put another way, the production of speech involves encoding operations and so one must expect that the inverse operations, like decoding in cryptography, will be inherently complex and liable to ambiguity.

An alternative view of speech perception does not link it to the motor system and so evades any need to run that machinery backward. It puts its dependence on auditory mechanisms, starting with feature detectors, and employs processes that are in principle describable by models and analytic rules, though these have yet to be discovered.

Clearly, the nature of speech perception is a central problem for speech research. Its relevance to computational linguistics, already discussed in connection with speech understanding systems, lies in how it affects one's choice of strategy in choosing methods for inferring the phonetic string from acoustic parameters, i.e., whether to stress analysis by synthesis with all its inherent difficulties or to concentrate on finding analytic rules, accepting the risk that they may not exist in any useful form.

Applications With Speech Involvement

One can be reasonably certain that the practical applications of automated language processing will not lag far behind the development of a technical capability.* It is easier to foresee examples that involve written-to-oral conversions than oral-to-written, so they will be discussed first.

Reading Machines

Synthetic speech as a reading service for the blind has already been discussed at length in terms of the research problems involved in setting up a central facility to make tape recordings from books. There is genuine need for such a service, especially for students and professionals. The practical objective is to have several service centers scattered across the country so that mailing delays will not be excessive. This will have to be preceded by a shakedown of the methods (still research oriented) and then some operating experience with a pilot center that uses production methods and equipment. It should be possible to accomplish both tasks within about five years, and to have begun the establishment of a network of service centers.

An obvious extension is to allow local users to have on-line access to the text reading facility and, as a second stage, to make this access available by telephone. The latter would pose formidable problems if character recognition were performed centrally and only text scanning were done remotely. It may be, however, that newer methods of feature extraction, or total recognition of the printed text, will have been developed by that time, and so would make the data transmission problem quite manageable. Nevertheless, real-time

*The current status of research and development in this area is reported in the Conference Record of the 1972 Conference on Speech Communication and Processing, April 24-26, 1972, at Newton, Mass. [The Record is available from the National Technical Information Service or the Defense Documentation Center, AD 742236.]

continuous processing poses very different problems from those of batch processing, some very similar to the problems encountered in real-time interaction with a speech understanding system.

Remote Retrieval of Information

The same technology that reads for the blind can be used to allow quick access to library holdings by telephone from a remote location. Many of the local requirements for such a service will be met for other reasons in any case, so the additional investment need not be large. Thus, some types of library holdings (abstracts, bibliographic information, etc.) are increasingly being supplied and stored on magnetic tape, with programs that provide fast access to desired items. With a little help from the reference librarian, machine-readable information could be found and processed by the library's computer to yield synthetic speech which the remote user could listen to by telephone. Such a service will not answer all needs, of course, but it should be valuable in many instances and it has the great virtues of requiring little by way of additional central facilities and of being able to use the existing telephone network instead of special terminals.

An obvious elaboration is to allow the human caller to speak directly with the computer---another application of speech understanding systems. Even without this complication, however, there are important linguistic problems involved in remote information retrieval. An obvious one is that much of the information now stored in machine-readable form is very "dense" in that it uses many abbreviations and graphic devices. Even connected text is often telegraphic in style. The reinterpretation of such information, now organized for the eye, to make it suitable for the ear becomes almost a condition for telephone access.

Computer Assisted Instruction (CAI)

Most CAI terminals operate solely with visual output and keyboard input, not because these modes are always optimal but because other modes pose major technical difficulties. Speech output, in particular, would be highly desirable in many cases, and is clearly the method of choice for much of the interaction with children in the lower grades. They could benefit from a great deal of content instruction if it were presented orally, but they do not have the reading skills to cope with it visually. For older students, too, oral information would often provide a useful supplement to visual displays.

This enhanced capability for CAI requires little more than adaptation of the text-to-speech techniques developed for the blind; in fact, the problem of providing good speech is easier, since the instructional text can be stored as a marked phonetic transcription that has been hand tailored to give natural sounding speech. Moreover, the storage requirements---hence, the possibilities for truly interactive CAI programming---are essentially the same as for literal text. Thus, the real utility of synthetic speech to CAI is likely to be far more dependent on imaginative programming than on technical limitations in providing spoken responses.

The ideal arrangement, in adding a speech capability to CAI, would be to let the machine respond appropriately to spoken responses by the student. Special purpose solutions, comparable to digit recognition, might work very well in many cases, especially with older students. But the greater need, and certainly the greater technical challenge, lies in making it possible for the younger student to interact in a reasonably free manner with his automated instructor. Clearly, this involves all of the problems of speech understanding systems as currently envisaged, compounded with the technical problems of processing children's speech and the linguistic problems of dealing with their free-form syntax. As a practical matter, it would be a mistake to hold back on the use of speech as an output in the hope of an early solution to the input problem, despite the many advantages that two-way speech would have in enlivening the interchange and removing artificial constraints on instructional programming.

Voice Typewriter

The prognosis for typing or typesetting under voice control is probably no better than that for voice input to CAI. It is apparent by hindsight that the choice of the voice typewriter as an initial target for research on speech recognition was a serious error. Such a machine must deal with unrestricted inputs and a wide range of speakers and dialects. One can scarcely imagine a practical task of greater difficulty! However, both the nature of the problem and paths to intermediate goals that might lead on to an eventual solution have become much clearer. Thus, on the one hand, what we know of the nature of speech tells us that pattern matching will never provide a general solution, no matter how sophisticated the techniques; on the other hand, the use of a "side in" approach to speech understanding problems of limited scope promises to be one of the paths that might eventually lead to general speech recognition.

Speech Understanding Systems

The nature of the problems, the difficulties to be expected, and some of the areas in which research in both computational linguistics and speech can be helpful have already been discussed. The tasks in which speech inputs are to be used were chosen as demonstrations rather than practical applications. Even so, they are difficult enough to pose very real research challenges.

The question is often raised about what, if any, really practical applications exist for voice input or, in more realistic form, what practical tasks there are that are not handled adequately by more conventional and less complex means. The ARPA Study Group listed some eight tasks as examples of practical applications: airline-guide information service, desk calculator (with voice input), air traffic controller, missile checkout (accepting spoken comments and questions from a human inspector), medical history taking, automatic protocol analysis, physical inventory taking (involving voice interaction with a human inspector), and robot management by voice. If none of these seems of compelling urgency, it may be in part a reflection of the fact that our capabilities in speech recognition are still so primitive as to shackle our imaginations.

Man-Machine Cooperation

In summary, there are several practical uses to which speech outputs from automated language processing can be put, and probably will be put in the near future, though the prospect for practical application of speech inputs seems more remote. But one is tempted to say about speech recognition that, like Everest, it is there---and eventually the challenge will be met.

Man-machine cooperation with computers is already a fact of life, though at present that cooperation can only be had on terms that are convenient to the computer. Again, one is tempted to say that such a state of affairs cannot continue; it is a safe prediction that man will insist on cooperation on his own terms. This means that computers must learn to listen as well as to talk. It will not matter much that this involves complexity and expense; if these were paramount considerations over the long term, we would all have telegraphs in our homes instead of telephones.

Instrumentation for Research and Training

The objective of the session's participants in their discussions of machines and speech was to consider not only promising areas for research but also needs and new possibilities for research tools. One suggestion that met general approval was that a state-of-the-art survey be commissioned to discover what the needs really are and to make generally available a knowledge of recent developments in the leading research laboratories. Very often, new devices or software which are built to fill a local need do not seem to the investigators to be sufficiently important to justify separate publication. Hence, they remain unknown, except to a handful of visitors.

This led to a discussion of how widespread the need might be for sophisticated new research instrumentation. The need is, of course, dependent on the number of centers in which basic speech phenomena are being studied intensively, and the prospect for additional centers. On this basis, instrumentation needs are comparatively modest quantitatively though crucial for the limited number of major research centers that do exist---perhaps half a dozen in the United States, and comparable numbers in Western Europe and in Japan. The establishment of additional centers is made difficult by the "critical mass" (for both men and equipment) that is needed to do effective research; indeed, the increasing complexity of adequate tools and the need for cross-disciplinary approaches seem likely to increase the pressures toward centralization of research. There are, on the other hand, both a need for well-trained people, and a number of good academic centers where training could be much improved by enough research equipment to make that training modern and realistic. The sound spectrograph is one such tool that is now rather widely available. Speech synthesizers, on the other hand, which could be at least as useful, are rather rare. This seems unfortunate, especially since the technology is well known and the costs are not excessive. Thus, a good background for many of the kinds of speech research described in preceding sections could be obtained with a mini-computer plus disc file, obtainable for $10-15,000. The point was made that the much larger computers already available for batch or time-shared use at many universities are not adequate substitutes for even a mini-computer that can be used on-line; in fact, very few computer systems can handle speech, primarily because of the high, continuous data rates that are required. A state-of-the-art review could be particularly useful to schools that wish to install a training facility of the kind described, not only in alerting them to the possibilities, but in providing detailed information that often takes a great deal of time to learn by trial and error---a familiar experience, summarized by one discussant: "the first program costs a year".

Some of the new developments and trends to be expected in research instrumentation will be cheap mass memories of very large size, and a new order of magnitude in central processor speeds. There was general consensus that the trend toward interactive systems that operate in real-time will continue, with a large payoff in research productivity. Likewise, new facilities for graphic output are becoming easily available and will be most useful.

In summary, although this part of the discussion found continuing progress and no urgent needs in the limited number of centers where most of the basic research on speech is done, it delineated a considerable need to upgrade awareness and training facilities in a much larger number of university centers in order to enable their graduates to become familiar with modern methods and problems. A state-of-the-art survey would be a useful initial step.

Conclusion

The group's discussions concerning Machines and Speech dealt mainly with the nature of the research problems that are encountered in incorporating speech into computational linguistics. Two specific applications---reading services for the blind and speech understanding systems ---were discussed at length. Both are examples of an increasing trend toward automated language processing and, in particular, extensions of this technology to the use of speech as an input-output modality.

The group identified a number of specific areas in which there is strong interaction between the usual domains of speech and computational linguistics. Thus, for example, the synthesis of natural sounding speech, starting with written text, requires information about the placement of sentence stress, about durations, and about pitch contours. This is information that is implicit in the structure of written sentences; hence, good synthesis would seem to require a capability for parsing. Conversely, in attempting to infer sentences in written form from an oral input, the suprasegmental information could provide much help in assigning structure to a string of phonetic elements. In general, many of the problems---and some very promising areas for research---lie in the conversions that must be made between language in written form and language in oral form. Thus, the addition of speech as input and output modes for automated language processing will necessarily focus attention on a whole set of problems that might otherwise pass unnoticed.

There are other areas of speech research that also interact with higher level processes. Thus, efforts to build detailed models of the processes of speech production (and to apply them to synthesis) must start with a description provided by phonology, and so cannot ignore the interface---presently missing---between speech processes and phonological rules.

Practical applications that make use of speech as an <u>output</u> from automated language processing are well on the way to being realized: reading services for the blind, remote retrieval of information by telephone, and a vocal response capability for computer assisted instruction. The prospects for processes that use speech as an <u>input</u> are more tenuous, though a major effort is under way to build demonstration models of speech understanding systems, i.e., computers that will accept instructions or questions via microphone. For the long term, there is little doubt that man-machine interaction will become increasingly important in a practical sense, or that there will be a steady pressure on the machine to conform to human convenience, i.e., to learn to talk and to listen.

The state of speech research here and abroad was also discussed briefly and it was noted that the trend toward concentration of research in only a few major centers is likely to continue because of the critica mass of men and instrumentation needed to deal with problems that are increasingly complex and multidisciplinary. But adequate research training need not be correspondingly concentrated; the provision of modest research facilities---in particular small computers used for synthesis studies---could do much to broaden the base of research training. A state-of-the-art survey and prospectus would be a useful first step.

LANGUAGE PERFORMANCE (PSYCHOLINGUISTICS AND DIALECTOLOGY)

John B. Carroll, Chairman

Although the scope of the topic assigned to this group was not defined in advance, the group took it to imply any application of computational techniques to problems of describing or modelling verbal output in actual communication situations. While such a scope could include the use of computational grammars and parsing techniques as discussed by other groups at the conference it was agreed that an emphasis would be placed on what kinds of information could be generated by the application of such techniques, assuming that the techniques had already been brought to some level of usefulness (rather than the specific problems of perfecting such systems).

The initial difficulty of setting the ground rules for discussion was further compounded by the fact that even the notion of "language performance" is not well defined, for in many ways it is related to the notion of "language competence", in the sense that the rules of "language performance" depend upon the rules of "language competence" and to a large extent may be identical.

Likewise, psycholinguistics is an ill-defined term. Some would apply the term to any scientific study of the use and understanding of language in language users; others would restrict it to the study of such matters as the special ways in which particular aspects of linguistic systems (phonology, syntax, etc.) are acquired and used.

It was recognized that computers are already very widely used in psychology, in the statistical analysis of psychological data, in the generation of stimulus arrays and experimental designs, and in "on-line" experiments in which subjects interact with computer-generated displays. (For references, see, for example, Cooley and Lohnes, 1971; Veldman, 1967; Tepas et al., 1972.) However, it is rare that any techniques similar to those of computational linguistics are involved in such uses; i.e., it is rare that rules relating to language structure are involved here. Nevertheless, it is difficult to draw the line; e.g., would a computer program for generating random strings of phonemes under certain linguistic constraints be regarded as falling under the heading of "computational linguistics"?

The topic of dialectology was treated rather summarily, almost as an afterthought, partly because most techniques thus far developed in computational dialectology are essentially "bookkeeping" techniques used to store and categorize data. However, several interesting ideas which were brought up by participants will be discussed in the report.

Under very loose ground rules that mere "bookkeeping" uses of computers would not be considered, the group proceeded to survey areas that had already received some study and/or that might hold promise for future work.

Studies of Linguistic Frequencies

Dangerously close to "bookkeeping" uses of computers, but still holding some possibilities of theoretical interest, are studies concerned with the distribution of linguistic elements with respect to frequency, i.e., what is the relative frequency of each of a given set of linguistic elements such as words, morphemes, phonemes, phoneme sequences, grammatical structures, etc.? Are there interesting ways in which to characterize these distributions? Can inferences be drawn from finite samples as to the characteristics of populations of infinite size? Work along these lines derives from that of Zipf (1935, 1949) and his predecessors (Estoup, 1916; Condon, 1928). Zipf studied various kinds of frequency distribution, frequencies of words, word lengths, Chinese characters, etc. and claimed that a certain law would characterize them, namely $fr = k$, where f is the frequency of a given word type, r is its rank (the number of items that occur f times or more), and k is a constant depending on the sample size. (This discussion is considerably oversimplified; for a more detailed discussion, see Miller & Chomsky, 1963, pp. 456-464.) Three major lines of further development are those of Mandelbrot (1961), who elaborates Zipf's function, Simon (1955), who generates a certain stochastic process, and Herdan (1960), Howes (1971), and Carroll (1969), who use a lognormal distribution or variants of it. In particular, Carroll and Howes have developed methods of estimating population distributions from sample distributions. The computations involved in all these applications are complex and require the use of high-speed computers. There still remain serious problems as to the generality and fit of the mathematical models that are claimed.

Computers are also used, of course, in the setting up of the raw frequency distributions from large corpora of data, but the sorting methods used here are simple and resemble those used in the development of concordances (Lamb & Gould, 1964). Among examples of extensive frequency counts done by computer are those of Kučera and Francis (1967), Carroll, Davies, & Richman (1971), and Roberts (1965). The first two of these counts were of the frequencies of words defined as graphic sequences; the Roberts count was of phonemes and their transitional probabilities, derived from a "hand-made" phonemicization of the word-types in a previously available corpus. One serious problem that has occurred in the graphic word count studies is that the results do not distinguish homographs, and fail to group the words by lemmas. Juilland and Chang-Rodriguez (1964) solved this problem by a hand post-editing of the computer output, but a better solution could be obtained if a highly accurate computer parsing procedure were available to be applied to the source corpus.

A related problem is that of deriving rules for English spelling-sound correspondences, either the rules for going from phonemes to spellings (Hanna _et al_., 1966) or the rules for going from spelling to phonemes (Venezky, 1970). Computers have been used extensively in this work, either for inducing rules from large corpora or pairs of printed words and their phonemicizations, or for testing rules on such corpora. Again, because of the anomalies of English orthography, the success of this work has thus far been limited by the difficulty of attaching appropriate semantic and grammatical readings for the words involved; presumably adequate parsing techniques would be of great help here. The practical importance of this work is that it would aid in the preparing of materials in the teaching of reading and spelling; further, it would supply links in any complete computerized system for going from speech to print or from print to speech.

Use of Computers in Describing and Analyzing Language Output

Many types of psychological investigations involve the eliciting or gathering of large quantities of verbal output. The sources and purposes are varied; examples are:

TAT (Thematic Apperception Test) protocols from individual subjects, to diagnose aspects of personality and motivation (Murray, 1938; McClelland _et al_., 1953).

Psychiatric interview data, to study themes in patient thought content or to gauge progress in therapy (Iker & Harway, 1965).

Folk tales from different cultures, to study common elements in thematic content.

Political speeches, propaganda, and the like, to assess political attitudes, subtle changes in propaganda "lines", etc.

Teacher-student interactions in the classroom, to study "classroom climate" and factors underlying effective teaching strategies (Bellack _et al_., 1966; Loflin _et al_., 1972).

In the past, verbal outputs of these kinds have been studied primarily by subjecting them to elaborate and time-consuming hand-coding schemes, but with the advent of the computer there has been increasing interest in automating these coding processes. One very general methodology that has been developed, based largely on computer techniques, is known as the General Inquirer program developed by Stone, et al. (1966). Essentially, this program is based upon the construction of specialized dictionaries wherein particular words are coded for specified themes or contents; the computer processes a text and counts the number of words associated with each of a number of themes as specified by the dictionary. In early forms of this program, words were counted simply as graphic shapes without regard to their actual meanings (which might count, for example, the word storm under the theme aggression when it referred to a weather phenomenon). Later versions of the program have included some attempts to disambiguate meanings in terms of context and grammatical form (Kelly, 1970), but it cannot be said that these attempts have utilized any truly sophisticated parsing rules.

A somewhat more sophisticated use of the computer to parse verbal output has been developed by Professor Baldwin of Cornell University (personal communication). In his studies of mother-child interactions, observers are taught a limited language for describing these interactions; it has been found possible to parse the resulting descriptions and thus obtain counts and other analyses of the voluminous interaction data thus obtained.

Barron (1970) developed a computer program for identifying and counting linguistic elements classified according to Fillmore's case grammar, and demonstrated sex differences in the verbal styles of men and women teachers.

Page and Paulus (1968) attempted to develop a computer program that would grade student-written essays as accurately as teachers do. Darnell (1970) developed a program for measuring the "clozentropy" of responses of foreign students learning English when they attempted to guess missing elements in English sentences.

This is only a very partial and selective listing of work in this general area. While all these programs have been partially successful, it is believed that work of this type could benefit very much from systems of language analysis that might be developed through a deeper understanding of language construction, i.e., the computer parsing techniques discussed by other groups at this conference.

Studies of Semantic Networks and the Subjective Organization of the Lexicon

One of the first major attempts to develop a quantitative approach to semantics was that of Osgood, Suci, and Tannenbaum (1957), embodied in the technique known as the semantic differential. The computational techniques were primarily those of factor analysis, applied to the very large data matrices of responses of subjects to semantic rating scales. A kind of zenith in this kind of work was reached in the atlas of semantic differential scalings published by Jenkins, Russell, & Suci (1958). Semantic differential data have to do, however, primarily with the connotative aspects of "meaning"; generally, this connotative space is found to contain three main dimensions, evaluation, activity, and potency. These dimensions probably have little to do with the semantic space of denotation. In more recent years, psychologists have attempted to determine the characteristics of denotational semantic spaces, largely using multidimensional scaling techniques. The work of Fillenbaum and Rapoport (1971) may be cited in this connection.

Other approaches to semantic analysis are represented by the work of Minsky (1968), Quillian (1969), Simmons et al. (1968), Reitman (1965), and others, all being concerned with how semantic information can be represented in computer memories in such a way that it can be accessed and manipulated in such "artificial intelligence" tasks as question-answering, determining semantic similarity, and simulating a dialogue. Psycholinguists have taken some interest in these formulations, but thus far the techniques they have employed to test hypotheses about semantic organization have utilized more conventional experimental settings (measurements of reaction time, etc.) rather than computer technology (e.g., Collins & Quillian, 1969; Myer, 1970). In fact, such techniques are probably highly appropriate for these studies, although one can also envisage a role for computer technology in interactive settings in which, for example, an individual's semantic networks are compared with those specified in a computer model.

Modelling of Speaker and Hearer Behavior

Psychologists have shown much interest in attempting to simulate human behavior, particularly higher mental processes involving language, concepts, and problem solving (e.g., Hunt, 1962; Reitman, 1965; Newell & Simon, 1972). Usually these attempts are formulated as computer programs. Simulation is viewed as essentially different from "artificial intelligence" in that the former is concerned with attempting to replicate, in the computer, the actual sequence of information processing activities that takes place in the human organism in learning a concept or solving a problem, whereas artificial intelligence is only concerned with procedures whereby the program can achieve, by any means found feasible, a result that resembles or even surpasses (in quality, speed, or accuracy) what could be achieved by human intelligence. There are serious problems in "validating" computer simulations because the actual processes being simulated are accessible to observation only indirectly, if at all. It can be proved mathematically that a given observed result can be achieved by a multiplicity (i.e., at least two) of different operational procedures; thus, the choice of which procedure is actually being followed by a given human subject on a given occasion is theoretically indeterminate. Nevertheless, persons working in the field of computer simulation believe that this type of work will help to narrow down the alternative possibilities.

The group discussed briefly the possibilities of simulating various types of language behavior. It was recognized that any such simulation would have to be based on an exceedingly well-organized and serious theory of language behavior, including not only rules for language competence but also for language performance. Nevertheless, the group was impressed with the advances that have been made with certain simple simulation programs such as that of Winograd (1971), wherein a robot accepts English sentences and performs corresponding actions in locating and picking up objects such as colored blocks. Kaplan's (1971) application of augmented transition networks to psycholinguistic investigation also seemed very promising.

Mention may be made of some partially successful attempts in this area. Uhr (1964) developed a program that could learn to "translate" from one language to another by comparing strings. Siklóssy (1971) has recently developed a program called ZBIE which (according to the author) "accepts tree structured descriptions of simple situations, and improves its capacity to express these situations in the natural language that it learns." Siklóssy gives examples of how either English or Russian sentences can be learned by the program in this way, and claims that his results "support the hypothesis that language-learning need not require postulated separate language-acquisition devices."

Ronald Kaplan reported unpublished work of Donald Olivier at Harvard dealing with the question of whether a computer program could "learn" to segment natural speech (e.g., into words) by use of frequencies of occurrence and transitional probabilities. Considerable success had been achieved by such a program, lending further support to the notion that the child as language learner needs only relatively limited acquisition devices.

Dialectology

In discussing possibilities of computer use in dialectology, the group noted that up to now, most work has concentrated on archiving information, computing indices of lexicographic similarity, and the like. The computer can indeed be of much use in such endeavors, but is largely merely a substitute for clerical work. Joseph Grimes reported that he had developed a program for accepting dialect information and plotting dialect areas. Some reports of the uses of computers in dialectology are to be found in Garvin (1965).

It was commented that in view of the large interest in American regional and minority dialects and the voluminous data that might be involved in any thorough study of such dialects, use of computers would be almost mandatory to achieve useful results efficiently.

Some mention was made of the possibility of using computers in glottochronological studies in which masses of data from different members of a language family would be analyzed for common elements. Some work of this sort has already been done (Carroll & Dyen, 1962). Mention was also made of the possibility of simulating language change, e.g., testing the operation of sound-law rules in both well-known (Indo-European) and less well-known language families.

Conclusion

It was felt that psychological and psycholinguistic studies could profit enormously from the theories and techniques now being developed in computational linguistics. Few psychologists are sufficiently aware of these possibilities.

References

Barron, N. Grammatical Case and Sex Role: Learning Differences in Interaction. Unpublished Ph.D. dissertation, University of Missouri-Columbia, 1970.

Bellack, A.A., H.M. Kliebard, R.T. Hyman, and F.L. Smith, Jr. The Language of the Classroom. New York: Teachers College Press, 1966.

Carroll, J.B. A Rationale for an Asymptotic Lognormal Form of Word-Frequency Distribution. (Research Bulletin 69-90.) Princeton, N.J.: Educational Testing Service, 1969.

Carroll, J.B., P. Davies, and B. Richman. The American Heritage Word Frequency Book. New York and Boston: American Heritage Publishing Co., and Houghton Mifflin, 1971.

Carroll, J.B., and I. Dyen. "High-Speed Computation of Lexico-Statistical Indices." Language 38(1962):274-278.

Collins, A.M., and M.R. Quillian. "Retrieval Time from Semantic Memory." Journal of Verbal Learning and Verbal Behavior 8(1969):240-247.

Condon, E.U. "Statistics of Vocabulary." Science 67(1928):300.

Cooley, W.W., and P.R. Lohnes. Multivariate Data Analysis. New York: John Wiley & Sons, 1971.

Darnell, D.K. "Clozentropy: A Procedure for Testing English Language Proficiency of Foreign Students." Speech Monographs 37(1970):36-46.

Estoup, J.B. Gammes Sténographiques. 2nd ed. Paris: 1916.

Fillenbaum, S., and A. Rapoport. Structures in the Subjective Lexicon. New York and London: Academic Press, 1971.

Garvin, P. ed. Linguistics and the Computer: A Casebook. Bloomington, Ind.: Indiana University Press, 1965.

Hanna, P.R., J.S. Hanna, R.E. Hodges, and E.H. Rudorf, Jr. Phoneme-Grapheme Correspondences as Cues to Spelling Improvement. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education (OE-32008), 1966.

Herdan, G. Type-Token Mathematics. The Hague: Mouton, 1960.

Howes, D. "Vocabulary Size Estimated From the Distribution of Word Frequencies." In: Helmer R. Myklebust, ed. Progress in Learning Disabilities, Vol.II, 94-110. New York: Grune and Stratton, 1971.

Hunt, E.B. Concept Learning: An Information Processing Problem. New York: Wiley, 1962.

Iker, H.P., and N.I. Harway. "A Computer Approach Towards the Analysis of Content." Behavioral Science 10(1965):173-182.

Jenkins, J.J., W.A. Russell, and G.J. Suci. "An Atlas of Semantic Profiles for 360 Words." American Journal of Psychology. 71(1958):688-699.

Juilland, A., and E. Chang-Rodríquez. Frequency Dictionary of Spanish Words. The Hague: Mouton, 1964.

Kaplan, R.M. Augmented Transition Networks as Psychological Models of Sentence Comprehension. Santa Monica, Calif.: Rand Corp., November 1971.

Kelly, E.F. A Dictionary-Based Approach to Lexical Disambiguation. Unpublished Ph.D. dissertation, Harvard University, June 1970.

Kruskal, Joseph B., Isidore Dyen, and Paul Black. "The Vocabulary Method of Reconstructing Language Trees: Innovations and Large-Scale Applications." In: F.R. Hodson, D.G. Kendall, and P. Tăutu, eds. Mathematics in the Archaeological and Historical Sciences, 361-380. Edinburgh: Edinburgh Univ. Press, 1971.

Kučera, H., and W.N. Francis. Computational Analysis of Present-Day American English. Providence, R.I.: Brown Univ. Press, 1967.

Lamb, S.M., and L. Gould. Concordances from Computers. Berkeley, Calif.: Mechanolinguistics Project, University of California, 1964.

Loflin, M.D., T.W. Guyette, N. Barron, and M. Marlin. "Reconstruction in the Analysis of Verbal Interaction." American Educational Research Journal 9(1972):101-112.

Mandelbrot, B. "On the Theory of Word Frequencies and on Related Markovian Models of Discourse." In: R. Jakobson, ed. Structure of Language in its Mathematical Aspect. Proceedings of the 12th Symposium in Applied Mathematics, 190-219. Providence, R.I.: American Mathematical Society, 1961.

McClelland, D.C., J.W. Atkinson, R.A. Clark, and E.L. Lowell. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

Miller, G.A., and N. Chomsky, "Finitary Models of Language Users." In: R.D. Luce, R.R. Bush, and E. Galanter, eds. Handbook of Mathematical Psychology, Vol. II, 419-491. New York: Wiley, 1963.

Minsky, M., ed. Semantic Information Processing. Cambridge, Mass., M.I.T. Press, 1968.

Murray, H.A., et al. Explorations in Personality: A Clinical and Experimental Study of Fifty Men of College Age. New York: Oxford, 1938.

Myer, D.E. "On the Representation and Retrieval of Stored Semantic Information." Cognitive Psychology 1(1970):242-300.

Newell, A., and H.A. Simon. Human Problem Solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Osgood, C.E., G. Suci, and P. Tannenbaum. The Measurement of Meaning. Urbana, Ill.: University of Illinois Press, 1957.

Page, E.B., and D.H. Paulus. The Analysis of Essays by Computer. Final Report. Storrs, Conn.: University of Connecticut, April 1968.

Quillian, M.R. "The Teachable Language Comprehender: A Simulation Program and Theory of Language." Communications of the ACM 12(1969):459-476.

Reitman, W.R. Cognition and Thought: An Information-Processing Approach. New York: Wiley, 1965.

Roberts, A.H. A Statistical Linguistic Analysis of American English. The Hague: Mouton & Co., 1965.

Siklóssy, L. "A Language-Learning Heuristic Program." Cognitive Psychology 2(1971):479-495.

Simmons R.F., J.F. Burger, and R.M. Schwarcz. A Computational Model of Verbal Understanding. Santa Monica, Calif.: Systems Development Corporation, April 1968.

Simon, H.A. "On a Class of Skew Distribution Functions." Biometrika 42(1955):425-440.

Stone, P.J., D.C. Dunphy, M.S. Smith, and D.M. Ogilvie. The General Inquirer: A Computer Approach to Content Analysis. Cambridge, Mass.: M.I.T. Press, 1966.

Tepas, D.I., et al. "Proceedings of the National Conference on the Use of On-Line Computers in Psychology." In: Behavior Research Methods and Instrumentation, 1972 (in press).

Uhr, L. "Pattern String Learning Programs." Behavioral Science 9(1964):258-270.

Veldman, D.J. Fortran Programming for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1967.

Venezky, R.L. The Structure of English Orthography. The Hague: Mouton, 1970.

Winograd, T. Procedures as a Representation for Understanding Natural Language. (MAC TR-84.) Cambridge, Mass.: Project MAC, Massachusetts Institute of Technology, February 1971. [M.I.T. Ph.D. dissertation, 1971.]

Zipf, G.K. The Psychobiology of Language. Boston: Houghton Mifflin, 1935.

Zipf, G.K. Human Behavior and the Principle of Least Effort. Cambridge, Mass.: Addison-Wesley, 1949.

SOCIAL IMPLICATIONS OF AUTOMATIC LANGUAGE PROCESSING

Donald E. Walker, Chairman

When computational linguistics first began to be identified as a field of endeavor about a quarter of a century ago, the social implications contemplated were only positive ones. Computers would process language atuomatically to provide mechanical translation (Weaver) and talking typewriters (Bush) or at least reactive ones (Mooers), and the consequences envisioned for society promised both international and trans-cultural cooperation and individual enhancement and augmentation. Recently, however, concerns about the impact of science on society have resulted in a sharp emphasis on the dangers that might result from computational linguistics. Invasion of privacy is envisioned, through listening in on telephone conversations or through the ability to address sophisticated queries to the increasingly comprehensive data files that are being accumulated. What seems needed at this time is a perspective both on the positive and on the negative implications of computational linguistics. What do we need to be concerned about? How can we guide the direction of the technological applications of our scientific advances? The discussions at the session on Social Implications of Automatic Language Processing were addressed to these questions.

There are two major contexts within which to examine the applicatior of computational linguistics. The first is provided by technological developments outside of the field, specifically, the proliferation and increasing availability of interactive terminals. Embedding similar devices in "the wired city" and allowing access to the capabilities of the computer through natural language over telephone lines or through radio and television links could have remarkable consequences for the life of the average citizen. The experiments in Reston, Virginia, make these possibilities more real.

The second major context for consideration is that of social needs. How does computational linguistics relate to contemporary urban problems or to the newly recognized responsibilities to people in isolated communities? Literacy and communication still are major problems. Social planning requires increasingly larger and more complex information bases for decision making. And there is increasing recognition of the right of the people both to know and not to be known without their knowledge. How and where does computational linguistics fit in?

It is reasonable before considering these questions directly to look more closely first at the scientific enterprise. And the initial topic for group discussion was in fact the sources and effects of the funding practices of the Federal Government. Quite apart from any specific information about technical content or projected goals, the identity of the agency supporting research can prompt attitudes toward the work. For example, some participants noted that money from the Department of Defense has been viewed as bad in principle, an interpretation complicated and confounded by the Mansfield amendment which has resulted in odd discrepancies between the descriptions of a project made by a university recipient and by the particular government agency sponsoring it.

One recommendation strongly urged was that all research funds (or at least those now handled by the Department of Defense) be provided through the National Science Foundation. Although the merits of such a proposal were recognized, it was argued that the Congress would never allocate as much money through NSF, so one consequence would be a drastic reduction in support for research. Equally compelling, however, was the argument that massive funding from a single source would inevitably narrow the range and variety of projects likely to be approved. Moreover, the other agencies need to have a knowledge of the basic sciences directly available in order to make informed decisions in support of development and engineering implementations over both long and short periods.

There is another effect of funding practices, less often recognized explicitly and with narrower "social" implications, specifically those for the particular scientific community. Massive concentrations of funding on a particular problem can have consequences for a field of inquiry. Support of mechanical translation in the late 1950's and early 1960's was presented as a case in point. Pressures toward a practical product in the absence of parallel support for the research essential to its achievement cannot succeed, and the consequences for computational linguistics because of the unreasonable expectations raised by the MT program were significant.

The Advanced Research Projects Agency's recent program to develop prototype speech understanding systems was given careful consideration by the participants for this reason. The assertion has been made that funding for this one activity is at least as large as the current support for all other speech research, if not larger. However, it was pointed out that the comparison may be misleading, because the ARPA program cuts across the full spectrum of language, including semantics, syntax, and morphology, as well as acoustics, phonetics, and phonology, and because it involves modeling the breadth of language capabilities. Nevertheless, the appropriateness of the concern was recognized; it is essential that some balance be kept, so that speech understanding can build on a continuing base of research.

One special problem needs to be kept in mind in this context. In the presence of large research programs, particularly attractive ones, there is a tendency (pronounced, if not always well articulated) for scientists to adjust their interests a little too much in accord with these programs. Consequently, funding agencies may find it necessary to overcompensate in making their decisions about projects proposed, deliberately supporting those that are not "in the mainstream".

The preceding remarks have involved narrow interpretations of the term social, although ones too often ignored in discussions of the implications of science. The major concern of the session certainly was with the effects of automatic language processing on the larger society. Consideration was given first to more positive implications.

Enhancement of communication still is expected to be a major consequence of successful work in automatic language processing. Mechanical translation can contribute to international communication or at least communication across national borders. There are functioning programs in MT, somewhat crude but both usable and used. The most interesting near-term prospect in this area is for computer aids to translation. Coupled closely with attempts to capture and model the translator's actions, this approach appears particularly promising.

Communication can also be considered with a more local focus. In the context of existing systems for personal file handling, the prospects for augmentation of human intellect (Engelbart) loom particularly large. Question-answering systems are laboratory exercises at this time. However, to the extent that they can be brought into interactive time-sharing systems, users who are specialists in technical areas that involve textual data should be much more efficient in their explorations through the existing literature and much better able to relate, edit, and even create documents in ways that add to the knowledge in their specialty.

The procedures that underlie the "augmentation of human intellect" of the research scientist can become part of the increasing automation in the library. The current experiments (e.g., Intrex) have made very little use of the potential of computational linguistics (and appropriately so). However, the critical barrier is the willingness of the library user to tolerate and, more importantly, to explore the potential of computer assistance. If bibliographic support proves satisfying in the library context, we should increase the sophistication of the procedures toward fact retrieval, with significant consequences for the effective dissemination of knowledge.

The same kinds of computer capabilities that are relevant for augmenting libraries and human intellects can be applied directly to education and to the important social issue of literacy which has such significant implications for participatory democracy. Few current computer-aided instruction techniques make use of computational linguistics. The possibility for meaningful and constructive dialogue with a student requires a degree of sophistication that can be provided only by techniques that penetrate beyond form into content. Moreover, such dialogues cannot be based on predetermined scenarios; they require the possibility of alternating questioning and answering phases from both student and program. Prerequisite for the development of any such capabilities is a comprehensive computational analysis of language.

From educational implications it is a natural extension to a consideration of the consequences of automatic language processing through interactive terminals in "the wired city". As remarked at the beginning of this section, through such a capability the citizen would have access to substantial computer power. Whether he is able to use that power depends critically on its accessibility through an interface that respects his customary modes of interaction. Techniques for automatic language processing must respect habits that are language-based, if systems are to be responsive to the general public.

The technical particularities of computational linguistics that would make for revolutions in education and in people's daily lives were not elaborated on in the discussions. Treated with somewhat more substance---but still with little in the way of specifics regarding the relevant hows of automatic language processing---were the implications for the medical and legal professions. Elicitation of medical histories, multiphasic health screening, the storage and retrieval of medical information, the more precise determination of variations in symptom patterns and diagnostic indices through processing large quantities of patient protocols, perhaps even diagnosis itself, all might be enhanced through automatic language processing capabilities. Comparable statements were made about the legal profession. Such enhancements certainly would make the practice of law and medicine more efficient and more effective and could be extended to other areas.

However, along with the positive social values associated with the potentials of computational linguistics and automatic language processing, a number of possibly negative consequences were considered. One of the most frequently cited relates to the current propensity in our society for aggregating information about individuals in centralized files. To the extent that natural language query capabilities are developed, these files could be accessed in violation of a person's privacy. Even if the retrieval strategies prove to be relatively unsophisticated, it is possible that personal files would be structured to reflect only a relatively small class of standard or stereotyped information items about an individual. The resulting impersonality might be valuable for statistical overviews, but it would blur meaningful distinctions among people.

The area of speech processing, recognition, and understanding again provides an illustration. Although voice authentication might prove valuable for the banking and credit card industries, it is viewed less sanguinely in relation to eavesdropping and personal surveillance. The possibility that "freedom of speech" could provide "instant evidence" looms as a serious threat to personal security, particularly with a technology whose accuracy leaves something to be desired. Wholesale recording of conversations is less attractive if it requires massive amounts of time to discriminate relevant from irrelevant segments. Screening techniques that just identified the presence of human voices would be helpful. If, further, it proved possible to identify particular human voices, and then to discriminate when they were talking about particular topics, the implications for the violation of human rights would become more and more important to consider.

Another kind of violence to human rights also needs to be considered, this one from the context of education. The use of crude and inadequate capabilities for automatic language processing in programs for computer-assisted instruction could lead to what one participant termed "robbing a generation of students of their intellect". It certainly seems that CAI has been relatively unproductive, but it has also been obvious, even to casual observers, that the particular programs used have been limited in scope. Introducing complex---but inadequate---linguistic interfaces could lead to expectations of competence that are not justified.

In the brief review of possible negative social consequences of automatic language processing, it was apparent that in some kinds of applications the dangers arise from success, in others from partial solutions applied prematurely to the scientific and technical problems involved. The participants also agreed that the professional involved in computational linguistics has a responsibility to illuminate both the problems and the potentials of his research.

Economic factors are relevant in these considerations with implications that may be variously positive or negative. Will automated language processing techniques increase or reduce the overall cost of computer-based facilities? If such systems are commercially feasible, will they be priced so that they are available only to particular socio-economic groups? Will they serve to increase or reduce conflict among these groups? These questions are not unique to computational linguistics, but they do merit careful examination.

Science exists in and necessarily has an effect on society. It is unlikely that a more perceptive view of the consequences of developments in atomic energy would have precluded the development of the atomic bomb. Similarly, work in so-called genetic engineering will be pursued in spite of the potentially disastrous consequences that have been projected. What seems essential is to demarcate clearly between what we know and what we do not know at any given time.

In recognition of the fact that social and political factors could coerce what might be considered premature use of automatic language processing technology, it was suggested that it might be necessary to increase support in certain areas so that distinctions can be made between science and "black magic". As a case in point, recent court decisions have sanctioned the use of "voiceprint" identification as evidence in legal proceedings. Since there is no consensus among specialists in phonetics and acoustics regarding the reliability of such identifications, it is essential that further research be conducted to establish the limits of credibility. However, one probable consequence of this research is the development of more precise techniques for the identification of specific individuals on the basis of voice characteristics.

There is no way of removing science from society. What is needed rather is to make more explicit and more accessible to the public the state of scientific knowledge at a given time. However, the area of computational linguistics has one characteristic that could make further developments society-serving in a particularly significant way. For a natural language interface to a computer can make data available to the public as well as to the government agency or the computer specialist. Further research certainly will result in simplifying access to computers by nonspecialists and in increasing the availability of computational capabilities. Social and political changes will be needed to allow for the most productive use of this accessibility. However, these changes do seem to be consonant with our country's movement toward a more open and more democratic society. Computational linguistics could contribute to our ability to treat each individual more uniquely, respecting more clearly his individuality. Accordingly, it does seem that the most significant social implication of automatic language processing can be that of providing "all computer power to the people". But computational linguists must make clear the dangers as well as the potential inherent in that vision.

PROFESSIONAL STANDARDS, ETHICS, AND EDUCATION

Sally Y. Sedelow, Chairman

During the sessions on professional standards, ethics, and education, the participants spent more time discussing the question of education for the computational linguist and seemed less concerned with other topics. This report reflects that emphasis.

Professional Standards

One of the problems that were brought up was the desirability of devising a system for evaluating or refereeing papers or technical reports which are sent around in mimeographed form. Such papers and reports sometimes create the illusion and, ultimately, reality of a sanctioned tradition, a canon, when it may be both undeserved and undesirable from the point of view of the profession as a whole. In the course of this discussion it was noted that computational linguistics is in a somewhat unusual position inasmuch as there is no single journal devoted exclusively to this area. Currently, computational linguists choose among journals such as Language, Computer Studies in the Humanities and Verbal Behavior, The Journal of Verbal Learning and Verbal Behavior, Computers and the Humanities, and Behavioral Science. It was noted that the Association for Computational Linguistics had looked into various publishing alternatives and no single satisfactory solution was found. On the other hand, Mechanical Translation, past publication of the Association for Computational Linguistics, folded up because it was not receiving a sufficient number of quality papers to sustain a regular schedule of publication.

One suggested resolution to the standards problem was the establishment of a panel of referees who would evaluate abstracts and articles which are distributed through informal channels. There was considerable objection to the prospect of evaluating abstracts but some positive response to evaluating articles.

The desirability of good program documentation and program-testing was also stressed, but no mechanism for assuring high standards was suggested.

Professional Ethics

Some professional associations, e.g., the American Psychological Association and the American Bar Association, have drawn up statements concerning ethics and standards. For example, one of the ethical issues with which both psychologists and lawyers are concerned is the preservation of confidentiality. Although some felt that a statement on ethics and standards for computational linguists might well be explored, nothing concrete as to either the content or the mechanism for drawing up such a statement emerged from the sessions.

Education

Most of the session participants' time and energy went into attempting to identify a core body of knowledge which a practicing computational linguist might ideally have. Participants noted that the areas listed need not imply courses, although courses might be a useful way to satisfy some or all of these needs. Nor, if courses were being considered, should each topic suggested necessarily imply as much as a semester or quarter term course. It did not seem fruitful to the participants who discussed this topic to try to specify programs appropriate for the B.A., M.A., and Ph.D. degrees, nor, in fact, to specify content of courses. Such prescription was avoided because variation among institutions is great as to academic areas (e.g., no linguistics department, no statistics department, etc.) and as to preparation of students deemed acceptable for any given degree program. For example, many linguistics departments assume that the student entering the M.A. program has had very little linguistics. Or, as another example, students with a background in computer science who desire to be computational linguists require supplementation in areas different from those required by, for example, someone from English. It is possible that in time, separate programs or departments in computational linguistics will emerge. But, for the present, it seemed most desirable and feasible to designate areas of knowledge in which a computational linguist should have competence. One other caveat is probably worth mentioning: that is, many currently practicing computational linguists do not have the range of knowledge represented by the areas listed below. Such a situation almost always exists when a new academic area, or emphasis, is emerging and the session's participants felt that academicians and scholars who see certain needs ---even those which they do not themselves fulfill---should, at this point, take comfort in their liberality of spirit and vision rather than feeling demeaned for falling short of the ideal. Actually such a situation exists in the long-established disciplines as well.

The core body of knowledge given below was compiled from informal suggestions and therefore should not be looked upon as exhaustive either in its contents or organization.

A. Formal Theory and Methods

Discrete mathematics (e.g., finite graph theory, combinatorics (elementary), semi-groups, Boolean algebra)

Logic (e.g., propositional and first order predicate calculus), naive set theory (together with a sample set of axioms)

Introductory Automata and Computability Theory

Formal Grammars and Languages

B. Computer Science

Computer programming as a skill, including at least one high level programming language. It is assumed that the introduction to computer programming also will provide an introduction to basic computer organization and to an awareness of, and experience with, a range of processing modes (e.g., interactive, batch, remote batch).

Data Structures and File Management

An introduction to the nature and range of programming languages, including some minimal experience with several programming languages.

Knowledge of how to construct as well as practice at constructing compilers

C. Linguistics

Phonology (stressing morphophonemics with less emphasis on phonetics and phonemics)

Syntax
Semantics
(morphology and discourse analysis, associated with both these areas, should be included)

Rule-writing techniques (tricks of the trade) for grammars

An understanding of the concept of modeling natural languages

Survey of theory available to the computational linguist

"Structure" course in some language (e.g., English or a non-Indo-European language; some participants felt that it must be a non-Indo-European language and others felt that as long as the course emphasized structure, even English---assuming the native speaker were English---would be acceptable)

D. Computational Linguistics

Knowledge of the state of the art (including heuristic/artificial intelligence approaches) with a thorough knowledge of three or four prototype systems

The areas specified above under A through D constitute the core recommended for a practicing computational linguist. In addition, some computational linguists might well choose to concentrate on a given area of specialization, implying, of course, additional knowledge.

For example, computational linguists specializing in speech must know experimental phonetics, experimental psychology, the calculus through differential equations, and signal processing.

Other area specializations might include:

Psycholinguistics
Sociolinguistics
Anthropological Linguistics
Historical Linguistics

Other computational linguists might well do ancillary work (this need not imply an emphasis, although it might) in such areas as:

Text Processing
Statistics and Probability
Lexicography
Computer Architecture
Computer Software Systems
Computer Graphics