AD-A041 288 IBM THOMAS J WATSON RESEARCH CENTER YORKTOWN HEIGHTS N Y F/G 5/8

A METHOD FOR STUDYING NATURAL LANGUAGE DIALOGUE.(U)

FEB 76 J C THOMAS N00014-72-C-0419

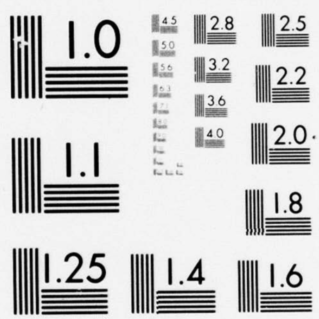
UNCLASSIFIED RC-5882 NL

| OF |

AD A041288

END

DATE FILMED 7-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADA041288

12

IBM Research

RC 5882
(#25427)
2/27/76
Computer
Science

Psychology

59 pages

A METHOD FOR STUDYING NATURAL LANGUAGE DIALOGUE

John C. Thomas

Behavioral Sciences Group
Computer Sciences Department
IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

This research was supported in part by the Engineering Psychology Programs, Office of Naval Research, Contract Number N00014-72-C-0419, Work Unit Number NR-197-020.

Yorktown Heights, New York

San Jose, California

Zurich, Switzerland

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: RC5882 (RC-5882)	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): A Method for Studying Natural Language Dialogue	5. TYPE OF REPORT & PERIOD COVERED: Interim Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): John C. Thomas	8. CONTRACT OR GRANT NUMBER(s): N00014-72-C-0419
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 197-020
11. CONTROLLING OFFICE NAME AND ADDRESS: International Business Machines, T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598	12. REPORT DATE: 27 Feb 76
	13. NUMBER OF PAGES: 62p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Code 455, Arlington, Virginia	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

349 250

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

dialog man-computer natural language metacomments business systems

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ABSTRACT: This report describes progress on an experimental technique for studying application - specific dialogues. In this technique, a user interacts via typed messages with a second person who is simulating a computerized natural language interface. The dialogues are all concerned with orderhandling and invoicing; however, they are collected in three different situations. The user is variously attempting to describe, understand, or diagnose an order-handling

(cont on p 1473B)

DD FORM 1 JAN 73 1473A EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(cont fr P1473A)

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

and invoicing system. Preliminary results indicate some of the ways in which the specific user task (pragmatics) influences dialogue features. The importance and complexity of having a natural language interface understand comments about the communicators and the communication itself as illustrated.

1473B

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

RC 5882
(#25427)
2/27/76
Computer
Science

Psychology

59 pages

A METHOD FOR STUDYING NATURAL LANGUAGE DIALOGUE

John C. Thomas

Behavioral Sciences Group
Computer Sciences Department
IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

ABSTRACT: This report describes progress on an experimental technique for studying application - specific dialogues. In this technique, a user interacts via typed messages with a second person who is simulating a computerized natural language interface. The dialogues are all concerned with order-handling and invoicing; however, they are collected in three different situations. The user is variously attempting to describe, understand, or diagnose an order-handling and invoicing system. Preliminary results indicate some of the ways in which the specific user task (pragmatics) influences dialogue features. The importance and complexity of having a natural language interface understand comments about the communicators and the communication itself is illustrated.

LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication elsewhere and has been issued as a Research Report for early dissemination of its contents. As a courtesy to the intended publisher, it should not be widely distributed until after the date of outside publication.

Copies may be requested from:
IBM Thomas J. Watson Research Center
Post Office Box 218
Yorktown Heights, New York 10598

ACKNOWLEDGEMENTS

I wish to thank Peter Sheridan, Ashook Malhortra, Irving Wladawsky and Lance Miller for their excellent comments on this manuscript. I also wish to thank Carol Coppola for her secretarial services and Donna Jones for her assistance.

I. INTRODUCTION.

This report describes progress on an experimental technique for studying application-specific dialogues in order to better understand how a computer might serve as a useful communication tool. In this technique, a subject interacts with a semi-automated dialogue system. This "semi-automated dialogue system" is actually another human who relies on questions and answers generated by computer insofar as possible. The procedure by which this human chooses questions and answers is being progressively automated. There exists a semantic network and a set of access routines to aid the human in retrieving relevant semantic information about the application. In the experimental dialogues, the subject engages in one of three tasks: 1. he attempts to understand how a particular order-handling and invoicing system works, 2. he diagnoses a problem with an order-handling and invoicing system or 3. he describes in detail a particular example of such a system. In any case, he accomplishes his task by carrying out a typed, interactive natural language dialogue with the semi-automated question-answering system.

This report does not present detailed analyses or conclusions based on these experimental dialogues. Rather, this paper is meant as a report of work in progress. One purpose of this paper is to put these experiments in perspective both with respect to other dialogue research, and with respect to the eventual construction of a fully automated, application-specific dialogue system. In Section II of this paper, a brief overview of the kinds of studies that have been done on dialogues will be presented. Naturally, that section is not meant as an exhaustive review but only as a framework within which to view the effort described in the remainder of the paper. In Section III, the purposes and strategies of the dialogue studies is outlined. A brief description is also given of the Mentor for Business Applications (MBA) project. Fuller descriptions of this system (Malhotra & Sheridan, In prep.) and the philosophy behind the system (Malhotra & Wladawsky, 1975) are available. Another purpose of this report is to describe in some detail the methods used in these experiments; this is done in Section IV. The final purpose of this report is to present some preliminary observations to demonstrate the kinds of data which are likely to come from these studies. This topic is covered in Section V.

II. REVIEW OF TYPES OF DIALOGUE RESEARCH

Most of the work on dialogues could be categorized into one of the following classes: 1) recordings of man-machine sessions 2) laboratory experiments 3) structural analyses. These types of studies are examined briefly below.

Recordings of man-machine sessions. Recordings of sessions between humans and current "natural language" computer systems are generally the province of computer scientists and computational linguists. Their method is to attempt to build systems that embody their proposed systems of analysis. While several interesting systems have been built and partial dialogues reported, (e.g. Weizenbaum, 1966; Winograd, 1972; Shank, Goldman, Rieger, & Riesbeck, 1973; Burton, 1974; Heidorn, 1974), it is generally difficult to determine from an examination of these reports which aspects of the system would be useful to someone designing a system for a slightly different purpose, or even for a system in a somewhat different domain. The portions of dialogues presented in such reports, moreover, are probably not representative of what may actually be encountered, but rather are chosen for illustrative purposes.

Laboratory experiments. Laboratory experiments on the effects of one or two manipulated variables on the effectiveness of comunication as measured by time or errors on some common task make up a second class of studies on dialogues. A favorite parameter here is the communications medium itself (e.g. Ochsman and Chapanis, 1974). Other studies have looked at such factors as the social desirability of the persons one has communicated with as correlated with the convergence of temporal speech rates (Natale, 1975). (Speakers who start conversing with widely discrepant speech rates show partial accomodation to each other's rate). Other investigators have examined the effects of various message, sender, and receiver characteristics on the effectiveness of a message in terms of changing beliefs and attitudes of the receivers of a communication. Other things being equal, beliefs and attitudes change more when the source of communication is perceived to be credible, attractive, or powerful. (See for example, pp. 528f Morgan and King, 1971) Credibility adheres not only generally to a source but is also affected by the content of the message relative to that source. Other studies report, for example, that a two-sided message (i.e., one that presents some arguements for both sides) generally produces more change than a one-sided message and that people who themselves have higher status are less likely to be influenced by a given piece of information. This type of study is usually but not exclusively performed by social psychologists. The results of such studies often are of practical importance in themselves; however, systems designers rarely use such information as much as they might. Knowing that the pattern of questions and answers that one finds in dialogues between humans depends on the perceived power relationships, for example, (See Mischler, 1975) could lead a system designer to consider how he wants his computer system to be viewed or to assess how in fact it is viewed by the user. However, the information obtained from such studies is certainly not of much help in determining the likely flow of content within a dialogue, or the types of syntactic structures that are likely to be used.

Structural analyses. Other studies concentrate on the structural analyses of certain aspects of dialogues. Here, the emphasis is on formal rules for describing some structural aspect of language, e.g. turn-taking (Sachs, Schegloff, and Jefferson, 1974). Unfortunately, since the reader is generally inadequately informed as to the semantics or the pragmatics that are relevant, the usefulness of such studies for those designing man-machine interfaces is limited. These structural studies have typically been carried out by linguists.

Empirical work on natural language systems. In addition to work in these three areas, other recent work is proceeding in other, less traditional, directions. Researchers at the Standford Research Institute (SRI) are conducting empirical studies to aid the development of a speech understanding system. (e.g. Walker, Paxton, Robinson, Hendrix, Deutsch, and Robinson, 1975) as are investigators at the University of Southern California (Balzer, 1974). At Bolt, Beranek, and Newman, studies are being carried out not only to develop sophisticated linguistic techniques, (e.g. Woods, 1973) but also to determine some of the mechanisms that humans use to make plausible inferences in order to answer questions (Collins, 1974). These research teams are combining broadly conceived observational studies of dialogue with building a natural language system. The dialogue research reported in the remainder of this paper most nearly fits into this last category.

III. PURPOSE AND STRATEGY OF THE DIALOGUE STUDIES.

Natural language It is a generally accepted belief in the Data Processing (DP) industry that hardware costs of computing will continue to decrease at the same time that software cost will continue to rise. (e.g., Boehm, 1973) In fact, the low cost of hardware could potentially enable many people to enjoy the utility of a computer, provided that there was a reasonably easy way for a person to instruct the computer to do useful things. Typical programming languages and operating systems provide many artificial barriers to the easy use of computers. One possible improvement is to provide formal interfaces that have minimized the unnecessary. The Query By Example system developed by Moshe Zloof at IBM research is a good example of this approach. (See Zloof, 1974 for a description and Thomas and Gould, 1974 for an evaluation). Another approach often recommended for making computers more accessible is to use natural language as the man-machine interface. (cf. Miller, 1974) Such an approach certainly offers a minimum of learning for the user, but is not totally without drawbacks. (See, for example, Hill, 1972 for a catalogue of potential difficulties.)

It seems likely that a major difficulty in developing new programs is in human-human communication. This view was supported by a Delphi survey that appeared in Datamation (Scott and Simmons, 1974). Again, the development of a natural language interface would allow people who wanted a job done to directly instruct the computer.

Working with the Mentor for Business Application project

There have been many projects aimed at the development of a natural language interface. Until very recently, most of these projects however, have concentrated upon the goal of making a system which could receive input or give output in syntactically correct sentences. This goal is necessary but hardly sufficient to the development of the computer as a useful communication tool. Little consideration has been given to psychological issues in communication. Such issues are vital, however, if the computer is to become a useful communication tool.

In fact, we all know that even when we are communicating with another human in a specified domain, there are vast individual differences in the effectiveness as well as the enjoyability of those communications. Some of the reasons for these differences clearly lie beyond syntax and domain-specific knowledge. A syntactic parser, knowledge of the world and powerful inference techniques are necessary but not sufficient capabilities for a useful natural language dialogue system. In fact, it has already been pointed out above that man-man communication in natural language is far from perfect. A natural language dialogue system should not only be capable of natural language dialogue but should use natural language expertly. Computers are not used for arithmetic calculations simply because they can perform additions that humans can perform but because computers can do so more quickly and accurately.

On the other hand, practical knowledge concerning what constitutes effective communication has grown little since the time of the Greek philosophers or before. The primary reason for this has been the complexity of natural language communication. This complexity has made analysis and experimentation very difficult. The development of computerized natural language systems however, makes possible complex analyses and even more importantly, on-line experimentation. The dialogue studies reported here were carried out in conjunction with the Mentor for Business Applications (MBA) project.

The goal of the Mentor for Business Applications project is to develop a system that would allow a businessman to describe, interrogate or diagnose an application in English. This project and the underlying philosophy have been described in detail elsewhere (Malhortra and Wladawsky, 1975, Malhortra and Sheridan, 1975). Working with such a group allows dialogue study to become more sophisticated as the results of previous observations and experiments are used to build more powerful natural language processing capabilities into the system. Focussing upon a particular world domain (in this case order-handling and invoicing systems) is also useful since it limits the content of the dialogues collected. In addition, the primary goals of the subjects are fairly well known. This does not make natural language dialogue analysis simple, but, in contrast to unconstrained dialogue or discourse, considerably reduces the uncertainty about crucial factors, such as the basic goals of the communicants.

Another advantage of working with the MBA group is that theoretical views of the communication process can be incorporated into newer versions of the automated dialogue system and thus tested out pragmatically by comparing dialogues employing a system that uses such theoretical constructs with one that does not.

Observation and intervention. The strategy for the initial phases of experimentation is to collect dialogues and observe. Later, once some analyses have led to a degree of understanding about dialogues, then it will be possible to intervene in the dialogues to test out the ideas generated through analysis. For example, it may turn out that rules may exist for predicting the kinds of statements that follow certain kinds of headers. (By "header" is meant an expression that indicates how to interpret the rest of the message. For example, the conjunction ",but" tells the receiver of a communication that a qualification or reservation will follow. Similar effects are discussed in more detail in Section V.) This header information might be crucial to the understanding of certain messages. The importance of headers could be tested by putting in a filter between the communicators that would eliminate the headers or substitute misleading headers.

Laboratory experiments

Once the analysis of complex dialogue situations leads to specific hypotheses about communication mechanisms, these hypotheses may be tested in more controlled experiments. For example, it is hypothesized that when one hears the connective word 'however' it has the effect of temporarily suppressing one's inference activity. This idea could be assessed by modification of Bransford and Frank's (1972) inference-testing technique. In this technique, subjects are given a series of English sentences. These sentences are connected and from them may be drawn reasonable inferences. Sometime after subjects read these sentences, they are given a recognition task in which they are to determine whether various sentences were contained in the set of sentences. Interestingly, subjects often report having seen sentences which were not actually present but only reasonable inferences from the sentences that were actually presented. This technique may prove sensitive enough to allow an examination of the effect of 'however' and similar "headers" upon inference.

IV. METHODOLOGY

Subjects Four types of subjects are being employed in these dialogue studies.

1. One group of subjects corresponds as closely as possible to the intended user population. These will be professional accountants, businessmen, and business analysts. Analysing the dialogues of these individuals will be most helpful in determining the actual kinds of interactions likely with a computerized applications dialogue system.

2. A second group of subjects will consist of professionals from other domains. These individuals will presumably have a variety of sophisticated strategies for finding out about complex systems, but not have any background knowledge concerning business and accounting procedures. Their dialogues will be of interest since they will contain

numerous instances of on-line learning. The eventual MBA system will need to be able to cope with situations in which the person using the system will need on-line tutorials conerning business concepts, though in actual practice this should happen infrequently. In order to collect an adequate number of instances of such on-line learning with experts, many more dialogues would need to be collected.

3. Also of interest will be dialogues with college students who have taken one course in accounting. The majority of psychological studies have used college students and it will be of interest to see if their behavior is qualitatively different from the other populations.

4. Finally, for a final contrast group, clerical workers with experience working within actual order handling and invoicing systems will be used. Presumably, such individuals will be familiar with the terminology of business and with a particular system, though their strategies for finding out about novel systems may be quite different from those of professional scientists and programmer- analysts.

<u>Tasks and Procedure</u> We are collecting dialogues in three situations concerned with order-handling and invoicing: system understanding, system diagnosis, and system description.

System understanding task. After subjects have had the basic purpose and procedure of the experiment explained to them, they are given three pages of typed instructions. These instructions explain to the subject that he will first be attempting to understand how a computerized order-handling and invoicing system is supposed to work and that his understanding will be tested by having him specify the output (invoices) that should result from particular orders. The instructions go on to explain that after the subject understands the system and fills out the invoices, that he will diagnose a difficulty with a particular firm's application program. (This program is supposed to conform to the system about which the subject has just learned). The instructions tell the subject how to use the IBM-3277 display scope, which is used as his input-output device. The instructions also include a brief verbal description of how the order-handling and invoicing system is supposed to work. Also given to the subject are a sample order, invoice, and customer master and item master files.

After the subject reads the instructions the experimenter asks whether there are any questions about the experimental procedures. The experimenter answers any questions about the experiment itself, but if the subject begins asking about the order-handling and invoicing system, the experimenter refers the subject to the 3277 display. The experimenter insures that the subject understands the keyboard of the 3277 and informs the person who will be communicating with the subject that the subject is ready to begin. At this point, the person who is simulating the system types a message to the subject telling him that he should begin asking questions. The subject waits until the word INPUT appears as a signal that the computer is ready to receive his messages to the system. Then the subject begins typing questions and comments to the semi-automated system in order to understand the order-handling and invoicing application. At some point, the subject feels that he understands the application and types a message to that effect to the semi-automated system. Then the subject is given a test of his understanding. He is given an order and a portion of the customer master file and the item master file. The subject must fill out specified fields in the invoice for the customer that sent in the order. If the subject fills this out correctly except for minor arithmetic errors, then he is ready to go on to the second part of the experiment. If his responses show that he has some remaining basic misunderstandings about the application, then he is requested to continue questioning until he is sure that he understands it, unless the experimenter judges that this subject will not be able to understand the system within reasonable time constraints.

System diagnosis. Subjects who pass the invoice generation test, are then given a diagnostic problem-solving task concerning a particular mythical customer's ('Magic Carpet, Inc.') program that was supposed to be an instantiation of the application he has just learned about. When the subject is ready for the diagnosis phase of the experiment, he is given a symptom of trouble. Then the subject's task is to determine the problem with the current order-handling and invoicing program. The subject does this by interacting with a semi-automated dialogue system in the same manner that he did during the system understanding phase of the experiment.

System description experiment. System description largely operates in the same manner as the system understanding experiment described above. In system description however, the subject is asked to choose one particular order-handling and invoicing system that he knows about and answer questions in terms of that one system. Subjects in this experiment must have considerable knowledge of business previous to the experiment. The subject is encouraged to draw flow diagrams and take notes to make sure that the system he will be describing is clear to him. Then the subject answers questions posed by the semi-automated dialogue system. In this case, the person with whom the subject is communicating is using an ACS questionnaire as a guide. The ACS questionnaire (1971), is a questionnaire that IBM customers fill out to aid salesmen in determining how the businessman currently does his applications so that the businessman will acquire appropriate software.

<u>Analysis</u> <u>techniques</u>. Collating the actual words used by the experts gives an indication of what would be necessary lexical entries in a completely automated dialogue system for the order-handling and invoicing domain. It is also important to determine the extent to which ambiguous words are disambiguated by knowing the domain of interest (in this case, order-handling and invoicing systems).

In addition to this lexical analysis, the sentences and phrases used by the subjects and by the human who simulates the system are being looked at from the standpoint of the necessary sytactic parsing programs. This is important, since ambiguities and constructions that are within the English language in the sense of competence, are not necessarily used very often, particularly when one limits the domain of discourse. Of particular psychological interest is the way in which subjects express ideas of logic, e.g., quantification and conditionals. These have been identified respectively as areas of difficulty in computer query languages (Thomas and Gould, 1974, Thomas, in prep.) and programming languages (Miller, 1974, Green, Sime and Fitter, 1975).

The dialogues are also being coded in terms of content. An attempt is being made to recode the propositions in the dialogue and the presuppositions into a network consistent with widely used psychological models of semantic memory. Particularly, the work of Frederickson (1975) and Norman and Rumelhart (1975), and Anderson and Bower (1973) is of use. An attempt is being made to encode the network in terms of certain primitive types of entitities and relations which are felt to be sufficient to describe all business concepts. The use of a limited number of primitives is a strategy to make inter-subject comparison easier and to increase aggreement between individuals attempting to recode sentences and phrases into a semantic network. The tentative categories of entities are Files, Documents, Processes, Systems, Terms, Policies, Messages, the Communication system (including both communicators and the line of communication between them). The primitive relations are purpose, definition, superset, structure, attibute-value (only applicable when more specific categories do not apply), succession, control, simulation, option, similarity (only used if specific comparison to another entity is made and no other category but attribute value applies).

V. PRELIMINARY OBSERVATIONS

To illustrate a few of the kinds of data that can be analysed from these dialogues, some preliminary observations are presented below. These are merely meant to be suggestive since the analysis techniques themselves are still evolving. In addition, only 12 sessions have been recorded so far. The behavior of the "semi-automated dialogue system" has also been changing during the course of these experiments. For all these reasons, these observations should only be considered suggestive. Nevertheless, it is felt worthwhile to present these observations since there are no existing records that the author was able to track down, which contained dialogues so close to the crucial issues of what is involved in computerizing business applications.

Vocabulary. The files for the dialogues were analysed to determine the frequencies with which various words were used. In these analyses, only exact tokens are counted together. (Thus, e.g., "file", "files", and "filed" are three separate words). Of the most frequent 100 words in the dialogues, 43 were among the 100 most commonly used words found by Carroll (1971). Table I contains the other 57 of the 100 most frequent words in the dialogue. In the third column are the frequencies as listed in the recent American Heritage frequency count edited by Carroll (1971). Also in this table, in the fourth column are the number of potential distinct meanings that these words have in English as listed in Webster's (1967) collegiate dictionary. When people have dialogues within a specific domain such as oder-handling, not only do the relative frequencies of words (i.e., lexical tokens) change, but there are also differences in the meanings that are common. For example, the token "order" has 14 distinct meanings listed in the collegiate dictionary. In the dialogues, 48 of the fifty occurences of this word were as a noun referring to set of records or as a noun adjective in a phrase like "order cards" or "order data". Once, order was used in the context: "In order to decide..." and once in the context: "In what order are...." Surprisingly, "order" (with no ending) was never used as a verb in the dialogues. The word "price" has 7 different meanings in the dictionary but again, it was only used in one way in the dialogues. The token "file" has 11 potential meanings and only one of these (noun for a set of records) was used in the dialogues. There were other additional words that were not terribly common in the dialogues but which were very related to the topics of conversation (as judged by the experimenter) and therefore should probably be lexical items understood by the semi-automated dialogue system. Examples of important content words used only once in the dialogues are "bill", "business", "charge", "equation", "financial", "grossprice3", "merchandise", "payment", "receipt". There were also a number of important function words that only appeared once each in the dialogues, e.g., "after", "again", "actually", "because", "cause", "causing", "necessary" and many others. Also of interest is the fact that there were many mispellings or typographical errors. Among the tokens that appeared only once in the dialogues, there are at least 61 mispellings or typographical errors.

One measure of the commonality of the words used in different dialogues was calculated according to the procedure outlined below. For each subject separately, the twenty-five words used most often in that dialogue were ranked according to number of occurences. The frequency with which a word on a given list also appeared on the lists for the other dialogues (maximum of 7) was determined. This was done for each of the 25 words in each of the eight lists. About 53 of the top 25 words were shared in common. The top 3,5, and 10 words were shared an average of 54, 54, and 56 percent respectively. Miller and Becker (1974) found that when subjects "Programming in natural English" they shared the top 3,5, and 25 words in common 71 percent, 62 percent and 44 percent respectively.

Syntax---conditional expressions

A more thorough look at the syntax is being made by another member of the MBA group, George Heidorn. A quick look at some of the logical connectives was made by the author, however, since the natural use of conditionals, disjunctives, conjunctives, and quantification is of interest psychologically as well as linguistically. The practical importance of these constructions for a computerized dialogue system is that linguists have shown a large number of possible ambiguities that spring from these constructions. A behavioral issue left open however, by the arguments of the linguists, is how often in actual dialogue these various possible meanings are actually intended or interpreted. Morever, though it is important for a complete theory of linguistics or cognition to be able to account for all these constructions, it is also of practical interest to have some idea how often the constructions are actually used. In contrast to programming languages, for example, some of which contain many conditionals, Miller and Becker (1974) found that when subjects were asked to specify procedures in natural language, they seldom used conditionals (e.g., "if red, then put in box.") Rather, non-programmers were spontaneously more likey to use qualificational statements such as "Put the red things in the box." In the dialogues collected so far, there are a great number of cases in which contingencies are expressed by qualificational statements. However, there are also a number of interesting cases of explicit conditionals. These appear to present a number of interesting interpretation problems for a natural language dialogue system. First of all, the pscyhological literature has made it clear for some time, that in fairly simplified laboratory tasks, subjects use 'if...then' constructions in a number of ways that are inconsistent with the logician's truth table for material implication. (See, for example, Staudenmeyer, 1973). In the complex tasks given to the subjects in the dialogue studies, again, several logical relations were referred to by 'if...then' expressions. On various occassions, 'if...then' constructions were apparently used to mean something akin to a logician's 'if', 'only if', and 'if and only if'.

In addition, there were a variety of surface structures that were used to express conditionality even in the relatively limited number of dialogues collected. At least one example of all of the expressions listed in Table I was found.

The referents of the conditionals that subjects used varied among the three tasks (describing an order-handling and invoicing system, understanding such a system, or diagnosing such a system). In the system description dialogues, most of the overt conditionals (that is, conditionals which used obvious logical constructions like "if", "only if" etc.) described business contingencies that were outside the scope of the computerizable aspects of business. In most cases, this was so clearly the case, that the subject probably was aware of this. The semi-automated dialogue system typically asked a question of the form "Do you want A or B?" or "Do you ever X?" The subject answered "yes" or "no" but went on to qualify or rationalize his answer as shown in the examples below. ('USER' here refers to the subject in the experiments.)

SYSTEM: May any item be backordered.

USER: Yes, as long as its available.

USER: No, our price reductions are normally a result of being overstocked or because of special purchases that we make.

USER: No. Substitution leads to customer dissatisfaction, shipping charges, and rehandling if returned....

In other cases, the contingencies expressed by the user referred to the intent or meaning of an earlier comment as illustrated below.

USER: If that is what you mean, then go with A.

USER: If you mean by special charges, fgt. and handling and sales, yes.

Still other contingencies referred to contingencies between capabilities of the system and the answers to questions asked by the system.

USER: ...However, if your system is capable of unit conversion, I would like to explore that area.

USER: Calculate by computer, if feasible.

During the system descriptions, there was no case in which an explicit contingency that should be executed in the program was communicated from the subject to the system. Generalization from this finding would be dangerous since this could be a function of the way in which the ACS questionairre is set up or a function of the particular subjects that were used. However, it seems clear that a natural language dialogue system that can help a user pick an instantiation of a program should be able to deal with conditionals of the form "If you mean X, my answer is Y" "If you can do X, I want Y." In addition, the dialogue system must make a strategic decision concerning what to do about conditional answers of the form "Yes, if X." where a direct computer test on X is impossible.

Every subject used conditional expressions at least once in the course of describing a system. In contrast, during system understanding, four of the eight subjects never used a conditional expression. Two of the subjects only used conditionals once each. Partly, this may be because the particular order-handling and invoicing system that the subjects were attempting to understand was conceived mainly in terms of data flow rather than control flow. Most of the conditional expressions that were used in system understanding were about aspects of the business that were outside the program per se.

USER: What does the system do if 'rdate' and/or 'recdate' is 'too old' and how is 'too old' determined? (This is done by another system, though there is no reasonable way that the user could have known this.)

In addition, there were some contingent questions of the form, "If X is Y (in the system) then tell me how you calculate...." There were no conditionalizations on an earlier meaning or upon the capabilities of the dialogue system, though the latter might change if users were using a real (i.e., fully automated) computerized system.

In contrast to the above two cases, during diagnostic problem solving every subject used conditional questions; and, with one exception, these were questions that asked about the actual contingencies within the program that did order-handling and invoicing.

The uses of quantifiers is described in much greater detail elsewhere (Thomas, in preparation). Briefly, in the system understanding dialogues there were only six instances in which subjects used words which could betoken explicit universal quantification ("all", "each", "every"). Only three of these could really be considered to betoken something like the logician's notion of universal quantification. In the system description task, there were many instances in which the experimenter asked questions with an explicit quantifier. Often a quantifier was also used in the answer. This may have been a strategy to eliminate the ambiguities that may arise in the use of English quantifiers.

Question types.

Each of the questions in the dialogue was classified according to the scheme outlined below; once for surface structure and once for the apparent intended meaning.

1. T-what questions. "What is an invoice?" Two important subclasses can be distinguised: "What is?" and "What kind of a?"

2. N-when questions. "When is the invoice generated?"

3. O-who questions. "Who sent this invoice?"

4. W-how questions. "How is tax computed?" Another type of how question with a slightly different tone means something like "How is that possible?"

5. H-which questions. "Which field in the item master file is for state tax?"

6. E-where questions. "Where is the control listing produced?" Other where questions essentially meant "whence?" and "whither?"

7. S-is/are or verification questions. "Is there more than one copy of the invoice summary produced?" These can also be with reference to capability, obligation, etc. Such verification questions can be expressed with various modals and state of being verbs. (e.g. should, can, does, will, did, may, do)

8. Y-why questions. "Why are some customers tax exempt?"

9. F-what if questions. "What if the warehouse personnel find an invalid item number?"

10. R-number questions. "How many invoice copies are produced?"

In addition, messages that are essentially questions can be stated as commands or statements. "Please define 'acrec'." "I am confused as to what 'acrec' means."

Of the questions that were asked during the system understanding phase, most (80 per cent) of the questions had easily recognizable surface structure cues that matched apparent intended meaning. In terms of surface structure, most of the questions were either "What" (30 per cent) "is/are" (22 percent) "how" (14 per cent) or commands (8 per cent). Each of the categories postulated above occurred at least once, however. There were 4 of the 125 questions that were not classified or included in the percentages above. Two of these questions were conditional questions. "Then does the system do anything with acrec information and if so, what?" Clearly this is just a conditional assemblage of a verification question and a "what"? question. Another similar example was "If grossprice 1 is the sum of the gextprices (not the extprices), then how do the discounts in the invoice column labelled 'epdpc' get included in netamt?" A similar but subtler example was of the form "Can you show me the backorder when the quantity..." Actually this is partly a question about the capability of the machine and, contingent on that capability, another question. The other question not included in the analysis was a user's "hello?" This could probably be rephrased as "are you there?" It probably should be thought of as a verification question about the communication system (rather than the order-handling and invoicing system). An automated application system may have to be able to deal with questions concerning its capabilities and about the dialogue itself as well as questions concerning the application.

"Headers."

Green, Sime, and Fitter (1975) point out that natural language sentences contain words like "and", "so", and "therefore" that convey structural or grammatical information. These words are referred to here as "headers". Thus, when a phrase is headed by "therefore" it signals that what follows is an inference. Balzer (1974) demonstrated that even when most of the specific content words are replaced in a passage by nonsense words, humans are able to infer much about the content, presumably relying largely on "headers".

These "headers" really contain information about the dialogue itself and hence can be considered as metacomments. Metacomments can explicitly refer to the communication process - e.g., "the following is an inference - A or B". Metacomments can also be implicit as in "therefore, A or B."

In the dialogues collected, there were only a relatively small number of basically different implicit "header" constructions. It seems as though writing a grammar for these would be fairly simple. In contrast, by making the dialogue itself or one of the communicators the explict topic of a sentence, a communicator may esentially use the entire English vocabulary rather than a clearly defined subset. While different businessmen may express business concepts in a fairly uniform manner, there is no reason to believe that this will be the case in expressing communication concepts. Consider the following variety of potential ways of expressing the notion that there is miscommunication.

"I am confused."

"I don't get what you mean."

"Come again?"

"How's that?"

"I seem to have a presupposition which isn't true."

"I think we're on different wavelengths."

"Our models of the world don't jibe."

"Help?"

"I think you need to educate me on that point."

"Please elucidate."

"Did you mean what I think you meant?"

This is not an unreasonable amount of variety. In fact, in the dialogues collected, while there were some implicit headers like "Now,", "however", and "but," that were used several times, there was almost no duplication in the explicit metacomments. To deal adequately with the variety of possible explicit references to the dialogue itself, an automatic application system would apparently need at least as extensive a network to deal with meaning, communication, confusion, etc. as to deal with business concepts themselves.

One may, however, legitimately ask the question whether it would be necessary for the system to deal with these explicit metacomments. Perhaps such comments concerning the dialogue itself could simply be ignored. However, the first example selection of dialogue in the section on miscommunication illustrates the possible disastrous effects of ignoring function words. In that case, the user and the system had been talking about two totally different kinds of discounts. After that discussion was over the system (in this case, recall, simulated by a real person) said "O.K. Do you also want a discount applied to the invoice total?" The user realized here that something was amiss. The user had thought all along that they were discussing a discount on the invoice total. If the system had merely asked "Do you want a discount applied to the invoice total?", the user could easily have thought that the system was merely reviewing their conversation. In this case, it was probably crucial that the system used the word "also". Conversely, if the user had described one kind of discount and the system had thought he meant something else, it would be crucial for the system to comprehend that the user was shifting to a new kind of discount when he said "I also want a discount applied to the invoice total."

<u>Miscommunication</u>. It has often been pointed out that the use of natural language does not eliminate problems of communication between humans and, therefore, it probably won't eliminate problems of communication between an

automated dialogue system and the human user either. In some of the dialogues, miscommunincation between the two humans appeared. Three examples of this are given below. The first example is from a dialogue in which an accountant is describing a particular order-handling and invoicing method to the semi-automated dialogue system.

> SYSTEM: For a given item, the extended item price is the product of that item's unit price by the quantitiy ordered. Do you discount from extended item price?
>
> USER: Yes, we give a 2 per cent discount if paid within 10 days.
>
> SYSTEM: So the extended price discount you offer does not vary either by customer or by item?

COMMENT: (At this point the user and the system are talking about two different kinds of discounts. The system, which is attempting to fill out the ACS questionnaire has as a referent a system which allows the price to depend upon the customer's price class code. In that system, the system first checks a customer's price class code and then uses that as an entry into a table of possible item prices. The user has in mind, however, the discount which is given to all customers if they pay the amount due within so many days.)

USER: Correct.

SYSTEM: Do all your customers enjoy extended price discount within the period of ten days?

USER: They all have the option, but not all take advantage of the discount.

SYSTEM: If I understand you correctly, extended item price discount is offered to all customers?

COMMENT: (note that this is the second time that the system refers to the discount as an item discount. However, this is not sufficient to clue the user that he is talking about a different type of discount.)

USER: Yes.

SYSTEM: Fine. Are all items eligible for the discount?

USER: Yes.

SYSTEM: O.K. Is a discount also applied to the invoice total amount (i.e. to the sum of all extended item prices)?

USER: There seems to be some confusion. The 2 percent discount is applied only to the total of the invoice.

SYSTEM: Then you do not apply a discount to each extended item price (i.e., to the total for each line item in the invoice.)?

USER: Correct.

As noted above, it was probably lucky in this case that the "system" used 'O.K.' and 'also' as clues that it was now switching to a different topic. Otherwise, the user might not have noticed the misunderstanding and his resulting application might have applied two separate 2 per cent discounts.

In the following example, an important issue is not clearly resolved from the dialogue.

SYSTEM: Do you wish any items ordered but not shipped listed by warehouse location or not?

USER: Yes.

SYSTEM: Fine. Do you wish any items ordered but not shipped listed on the customer's invoice?

USER: Yes, and I would like a backorder file.

SYSTEM: Good. If I understand you correctly, when inventory quantity of an ordered item is insufficient to fill the order for that item, you backorder the item and make no attmept to substitute another item for the missing ones. Is that correct?

USER: Yes, substitutions should be made only by the selling agent.

SYSTEM: Right. Regarding your invoice numbering system...

COMMENT: What is not clear from the transaction above is what the user wants done in cases in which there is sufficient quantity on hand to partially fill a customer order. It is clear that a backorder should be produced. But does the user want: 1. A partial fulfillment of the order sent to the customer immediately and the remainder upon receipt of more inventory or 2. The entire order to wait until there is sufficient quantity on hand for filling the entire order. Backordering is mentioned again in this dialogue but that particular ambiguity is never resolved.

In the following example of a system understanding dialogue, the user is not familiar with business concepts.

USER: What do you mean by "heading", "body" and "totals"?

SYSTEM: The heading contains customer-related information, the body item related information, and the totals information and calculations related to the total amount due from customer.

USER: What items are related to the customer, what items are related to the body?

COMMENT: Direct questioning of the user by the experimenter made it clear that the word "item" as used in her questions could have been replaced by "thing". In other words, she was not referring to the objects sold by Magic Carpet, but using "item" to refer to units of *information*. Furthermore, the user did not see any potential ambiguity in her question. Conversely, the human at the other end took "item" to mean articles sold by Magic Carpet and answered the question accordingly.

SYSTEM: The *invoice body mentions* those, and only those, items ordered by the customer.

Thus, even context will not allow a system to choose the intended meaning of a word in all cases. Note that although neither communicator detected the ambiguity about "item", the system still gave an informative and useful response to the user.

Subject Differences. The MBA system is being designed for people who understand an application area. Because those who do not understand the application area (at least in the same terms that the system does) may nonetheless occassionally interact with the system, the MBA system should at least have the facility to recognize when the person it is interacting with is not sufficiently familiar with the business domain to engage in a fruitful dialogue. At that point, the system might attempt a tutorial or politely excuse itself. The tremendous differences between individuals familiar with the application domain and those not famliar is illustrated by the following three dialogue fragments. In these cases, each subject's task was the same ---to understand the order-handling and invoicing system of Magic Carpets, Inc., by asking a series of questions about the system. The first subject had no previous business experience or knowledge of computers. The second dialogue fragment is from a subject who had considerable experience with computers but no previous knowledge of business terminology or how an order handing and invoicing system worked. The third fragment is the beginning of a dialogue with an accountant quite familiar with most phases of business accounting systems.

SYSTEM: Hello there. Do you have any questions?

USER: Who sent this invoice?

COMMENT: already the user seems confused in some basic way. "This invoice" seems to imply that this user is interpreting the written description of the computerized order-handling and invoicing system and accompanying example documents as the exact things that he is to learn about rather than as examples to illustrate how a general system works.

SYSTEM: Let me clarify. Orders are sent to the firm, invoices are generated, then sent to the customer.

USER: What kind of customer do you handle?

SYSTEM: Customers interested in ordering rugs and carpets.

USER: Very funny.

SYSTEM: Please clarify your last remark.

The following exchange is illustrative of the middle part of the same dialogue.

SYSTEM. The invoice consists of a heading, body, and totals.

USER: What do you mean by "heading", "body", and "totals".

SYSTEM: The heading contains customer-related information, the body, item-related information, and the totals information and calculations related to total amount due from the customer.

USER: What items are related to the customer, what items are related to the body?

SYSTEM: The invoice body mentions those, and only those, items ordered by the customer.

USER: What do you mean by item related information?

COMMENT: Note that it would probably be quite frustrating for the user to receive the previous answer again. For this reason it would probably be wise for an interactive system to keep some track of its own answers as well as the user's previous questions.

SYSTEM: item-related information includes the following: item number, item description, quantity ordered, quantity shipped, unit price, gross extended price, extended price discount, net extended price, federal state, and local taxes, extended weight.

USER: What is a gross extended price, what is an extended price discount, net extended price, and extended weight?

SYSTEM: gross extended price, for a given item, is the product of quantity shipped by unit price. Extended price discount is the percentage discount from gross extended price permitted a given customer on any item ordered. Extended weight of an item is the product of quantity shipped by unit weight.

USER: Help!

SYSTEM: Which notion is causing difficulty?

USER: All notions!

(Note the futility of a natural language processor interpreting this "all" as equivalent to the logician's concept of universal quantification!)

It seems clear from this dialogue that a user completely unfamiliar with business terminology may quickly be overwhelmed by a barrage of technical terms. In the next example dialogue, the user is a psychologist and systems programmer who also does not understand business terminology. Nevertheless, the type of questions asked are quite different.

SYSTEM: Hello. Please fire away when ready.

USER: Please define what a 'control listing' is.

SYSTEM. A control listing is a listing of input data to verify its correctness.

USER: Define a 'customer master file'.

SYSTEM: A customer master file is an ordered collection of records containing permanent information about customers.

USER: Describe the contents of each record. (Though this user seems more analytic than the previous one, this does not prevent difficulties in communication as illustrated below).

USER: (After system supplied the list of fields in the customer master file.) Is the 'active record code' pertinent to the inventory management system or is it simply used by the data management system?

SYSTEM: The active record code indicates the frequency of reference to each record.

USER: Answer my previous two questions with respect to the 'item master files.'

(Note here that the user does not actually mean his immediately preceeding questions but rather is referring to the two questions before the one concerning 'active record code'. Somehow the dialogue system would have to guess that the user is referring back to analogous questions rather than taking the number designation of the question too literally.

Even later in the dialogue it is clear that this user still lacks a clear understadning of some of the business concepts. For example, the user asks: "Is the access to the master files contingent on (that is, is it performed during) the daily update of the 'control listing', or is the master file access a separate (asynchronous) occurrence?" Note that a control listing in the system under discussion is not updated at all; rather a new control listing is produced each day. It is also interesting to note that many of this programmer's questions seemed to presuppose that the system was a control flow system.

The following sections of dialogue between an expert accountant and the semi-automated dialogue system show clearly that the expert was quite conversant with the terminology.

SYSTEM: (The system is ready for your first question.)

USER: Do you have any invalid or incorrect arithmetic calculations in the current program.

SYSTEM: no.

USER: On Cr. items, how are these generated and how long are these held in memory before a cr. invoice is generated without the customer filing another order.

SYSTEM: By 'cr items' do you mean 'credit items'., and are these the same as returned items?

USER: Yes.

SYSTEM: Since we deal with regular customers, we give them credit on the next order.

USER: When the computer indicates an invalid item or an invalid customer even though the order is manually screened first, what happens?

During the problem diagnosis phase of the experiment, the user was equally adept.

SYSTEM: (You are now ready for the problem diagnosis phase. Please ask your first question.)

USER: It seems that the backorder qty for backorders is being picked up from the wrong field or that the calculation for b.o. items is not working or missing from the system. Can either of these be the problem?

SYSTEM: Yes.

USER: is the calculation being made.

SYSTEM: If you mean, calculation of backorder quantity, the answer is yes.

USER: Right. then I suggest you have your programmer check the file id. he is picking up for b.o. item qty. I strongly feel order qty. is being picked up.

SYSTEM: You're absolutely correct. (Thank you for your cooperation. I hope you enjoyed the session.)

<u>Summary</u>

The results so far must be considered as preliminary.

Nevertheless, they are instructive in several ways. First, it seems clear that a natural language interface must be able to deal with at least some metacomments (e.g., "However,...", "I didn't understand your last question.", "Now,...") Second, the way in which various expressions (e.g., conditionals) are used is heavily dependent upon pragmatics of the dialogue, not just the semantics. Third, users of different backgrounds will interact quite differently with a natural language interface--so differently, in fact, that the interface should probably be capable of discriminating professionals from nonprofessionals and take contingent appropriate action.

Clearly, the many speculations put forth here need to be tested empirically. Some of these speculations could be best investigated within the framework of experimental dialogues. These could be collected when a user interacts with a more sophisticated semi-automated system with modular capabilities. Other speculations should be tested in separate laboratory experiments (e.g., the notion that certain "headers" like "therefore" partly function as commands to a receiver's operating system.

Comments and criticisms regarding these ideas are most welcome.

REFERENCES

Anderson, J. R., & Bower, G. H. Human Associative Memory. New York: John Wiley, 1973.

Application Customizer Service System/3 (Disk) Sales & Distribution Questionnaire Publication No. S320-1165-1 IBM Corporation. (Sept 1971)

Balzer, R. M. Human use of world knowledge. Univ. S. Calif. Report ISI/RR-73-7, March, 1974.

Boehm, B.W., Software and Its Impact: A Quantitative Assessment. Datamation, May, 1973, 48-59.

Bransford, J.D., and Franks, J.J., The Abstraction of linguistic ideas: A review, Cognition, 1972, 1 (2/3), 211-249.

Burton, R. R. A Semantically Centered Parsing System for Mixed-Initiative CAI Systems. Paper presented at the Association for Computational Linguistics Conference, Amherst, Mass., July 1974

Carroll, J. B., Davies, P. & Richman, B. The American Heritage Word Frequency Book, Boston: Houghton Mifflin Company (1971).

Collins, A., Reasoning From Incomplete Knowledge, paper presented at the Fifteenth Annual Meeting of the Psychonomic Society. Boston, Mass, Nov., 1974.

Fredericksen, C. H. Representing Logical and Semantic Structure of Knowledge Acquired form Discourse. Cognitive Psychology, 1975, 7, 371-458.

Green T.R.G., Sime, M.E., and Fitter, M., Behavioral Experiments On Programming Languages: Some Methodological Considerations. MRC Social and Applied Psychology Unit, Memo No. 66; The University, Sheffield; 1975.

Heidorn G. E. English As A Very High Level Language For Simulation Programming. Proceedings of A Symposium on Very High Level Languages. March 28-29, 1974

Hill, I.D., Or Would It?, The Computer Bulletin, 1972, 16, (6).

Malhotra, A. & Sheridan, P. Experimental Determination of Design System Requirements for a Program Explanation System. IBM technical report (In preparation).

Malhotra, A. & Wladawsky, I. The Utility of Natural Language Systems. IBM technical report (RC-5739.).

Morgan, C. T. and King, R. A. Introduction to Psychology, New York: McGraw-Hill, 1971

Miller, L.A., Programming by Non-programmers, International Journal of Man-Machine Studies 1974, 6, 237-260.

Miller, L.A., and Becker, C.A., Programming in Natural English. IBM Technical Report, RC 5137, November 1974

Mishler, E. G. Studies in Dialogue and Discourse: II. Types of Discourse Initiated by and sustained through questioning. Journal of Psycholinguistic Research, 1975, 4 (2) 99

Natale, M. Social Desirability as related to covergence of temporal speech patterns.
Perceptual and Motor Skills, 1975, 40, 827-830.

Norman, D. A., & Rumelhart, D. E. Explorations in Cognition, San Francisco: W.H. Freeman, 1975.

Ochsman, R. B. & Chapanis, A. The Effects of 10 Communication Modes on The Behavior of Teams During Co-operative Problem-solving. Int. J. Man-Machine Studies,

Sachs, H., Schelgloff, E. A., & Jefferson, G. A simplest systematics for the organization of turn-taking for conversation, Language, 1974, 50, 696-735.

Scott, D.F. & Simmons, D.B. Programmer productivity and the Delphi technique, Datamation 1974 (20) (5),71-73

Shank, Goldman, Reiger, and Riesback, 1973, Margie: Memory, Analysis, Response Generation and Inference on English, Third International Conference on Artificial Intelligence, 1973, 255-261.

Skinner, B.F., 1957, Verbal Behavior, New York: Appleton-Century-Crafts, 1975.

Studenmeyer, H., Understanding Reasoning. Unpublished Ph.D. Thesis, University of Colorado, 1973.

Thomas, J.C., and Gould, J.D., A Psychological Study of Query by Example, IBM Technical Report RC-5124, November, 1974.

Walker,D.E., Paxton,W.H., Robinson,J.J., Hendrix,G.G., Deutsch,B.G., and Robinson,A.E., 1975, Speech Understanding Research, SRI Annual Technical Report, June, 1975.

Webster's Seventh New Collegiate Dictionary, Springfield Massachusetts: G. and C. Merriam Company, 1967.

Weizenbaum, J., 1966, Eliza - A Computer Program For the Study of Natural Language Communication Between Man and Machine, _Computational Linguistics_, Volumn 9/ Number 1/ January, 1966.

Winograd, T.H., 1972, A Program for Understanding Natural Lanugage, _Cognitive Psychology_, 1972, _3_.

Zloof, M.M., Query by Example, IBM Technical Report RC-4917, July, 1974.

TABLE 1

Generally uncommon words that were frequent in the dialogues.

DICT.	DIALOGUES	AMER. HER.	MEANINGS
files	.0018	.0000033	11
taxes	.0018	.000020	2
percentages	.0020	.0000011	4
production	.0020	.000044	3
sent	.0020	.00016	1
charges	.0022	.000019	14
listing	.0022	.0000068	13
shipped	.0022	.000020	14
active	.0024	.000034	12
amount	.0024	.00013	5
group	.0024	.00031	8
local	.0024	.000056	5
problem	.0024	.00022	4
control	.0025	.00011	5
during	.0025	.00037	2
field	.0025	.00018	10
sum	.0025	.00016	7
federal	.0027	.000053	6
information	.0027	.00016	5
last	.0027	.00059	10

applicable	.0029	.00000058	1
txcd2	.0029		0
chain	.0031	.000041	3
customers	.0031	.000014	2
date	.0031	.000041	12
indicates	.0031	.000031	2
yes	.0033	.00026	4
state	.0036	.00028	9
percentage	.0038	.0000090	4
question	.0038	.00017	5
backorder	.0040	.00000	0
ordered	.0040	.000057	1
unit	.0043	.00020	2
items	.0045	.000038	5
ready	.0051	.00023	3
system	.0051	.00022	4
distant	.0054	.00000098	5
please	.0056	.000099	6
quantity	.0058	.00000078	4
gross	.0060	.0000062	9
code	.0065	.000030	3
record	.0069	.000139	7
extended	.0071	.000022	3
invoice	.0071	.00013	4
master	.0078	.000080	7
order	.0090	.00029	14
file	.0092	.000020	11

tax	.0107	.000046	2
price	.0134	.000063	7
item	.0163	.000024	5
customer	.0184	.000011	2

TABLE II

Cumulative proportion of total tokens accounted for by the most common types.

American Heritage word count (1971)	System Understanding Dialogues	No. of different types
.0731	.0827	1
.1016	.1243	2
.1278	.1463	3
.1522	.1668	4
.1759	.1868	5
.1952	.2064	6
.2069	.2238	7
.2164	.2393	8
.2257	.2536	9
.2349	.2671	10
.2722	.3218	15
.2996	.3615	20
.3409	.4261	30
.3657	.4775	40
.3949	.5177	50
.4321	.5782	70
.4685	.6490	100

.5117	.7240	150
.5298	.7588	180
.5398	.7783	200
.5491	.8172	250
.5788	.8420	300
.6178	.9130	461

Table III

(Capital letters replace various technical business words here since the
emphasis is meant to be on the form of the expressions.)

"...as long as..."

"If your system...I would"

"No, but ...,if needed."

"In some cases, e.g. X, ... In this case, ..."

"A. Unless X. If X then B."

"Is X contingent on Y?"

"Then does the system...?"

"Suppose that X were Y. What would Z be?"

"Only if we are X."

"No, our X are a result of Y or because of Z."

"This will work if X. If this is not the case, then..."

"Then 1 =] tax and 0 =] no tax?"

Some of the contingencies expressed in "natural language" were fairly
complex.

"I am trying to determine A. I notice that X in K could be Y with the Z in L if the H in K were also H in the L;

"Yes. This can occur when selling to both wholesale and retail customers."

"Does the system decide to X only if both Y and Z?" One subject expressed a contingency in what he referred to as "bastardized algol" and he expressed another contingency in terms of a simple Boolean expression.

TECHNICAL REPORTS DISTRIBUTION LIST

CODE 455

Director, Engineering Psychology (5 cys)
  Programs, Code 455
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Defense Documentation Center (12 cys)
Cameron Station
Alexandria, Virginia 22314

Director, ONR Branch Office
ATTN: Dr. J. Lester
495 Summer Street
Boston, Massachusetts 02210

Director, ONR Branch Office
ATTN: Dr. M. Bertin
536 S. Clark Street
Chicago, Illinois 60605

Director, ONR Branch Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, California 91106

Director, ONR Branch Office
ATTN: Mr. R. Lawson
1030 East Green Street
Pasadena, California 91106

Dir., Naval Research Laboratory (6 cys)
Technical Information Division
Code 2027
Washington, D. C. 20375

Dir., Naval Research Laboratory (6 cys)
ATTN: Library, Code 2029 (ONRL)
Washington, D. C. 20375

Mr. John Hill
Naval Research Laboratory
Code 5634
Washington, D. C. 20375

Office of Naval Research
Information Systems Program, Code 437
800 North Quincy Street
Arlington, Virginia 22217

Office of Naval Research
Operations Research Program, Code 434
800 North Quincy Street
Arlington, Virginia 22217

Office of Naval Research
Naval Analysis Programs, Code 431
800 North Quincy Street
Arlington, Virginia 22217

Office of Naval Research
ATTN: Dr. K. T. Wallenius, Code 436
800 North Quincy Street
Arlington, Virginia 22217

Office of the Chief of Naval
 Operations, Op-987E
Department of the Navy
Washington, D. C. 20350

CDR H. J. Connery
Office of the Chief of Naval Operations,
  Op-987M4
Department of the Navy
Washington, D. C. 20350

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code AX
Washington, D. C. 20380

Dr. Heber G. Moore
Hqs. Naval Material Command
Code 0331
Department of the Navy
Washington, D. C 20360

Mr. Arnold Rubinstein
Naval Material Command, NAVMAT 03424
Department of the Navy
Washington, D. C. 20360

Commander, Naval Electronics
  Systems Command
Command and Control Div., Code 530
Washington, D. C. 20360

Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, D. C. 20360

Bureau of Medicine and Surgery
Human Effectiveness Branch, Code 713
Department of the Navy
Washington, D. C. 20372

CDR Robert Wherry
Human Factors Engineering Branch
Crew Systems Department
Naval Air Development Center
Johnsville
Warminster, Pennsylvania 18974

LCDR Robert Kennedy
Human Factors Engineering Br., Code 5342
U.S. Naval Missile Center
Point Mugu, California 93042

Lt. Col., Henry L. Taylor, USAF
OAD*E&LS) ODDR&E
Pentagon, Rm. 3D129
Washington, D. C. 20301

Mr. Richard Coburn
Head, Human Factors Division
Naval Electronics Laboratory Center
San Diego, California 92152

Dean of Research Administration
Naval Postgraduate School
Monterey, California 93940

Navy Personnel Research and
Development Center (Code 10) (5 cys)
San Diego, California 92152

Mr. James L. Long
Weapons Systems Research (N-332)
Naval Education and Training Command
Naval Air Station
Pensacola, Florida 32408

Human Factors Dept., Code N215
Naval Training Equipment Center
Orlando, Florida 32813

Dr. George Moeller
Head, Human Factors Engineering Branch
Submarine Medical Research Laboratory
Naval Submarine Base
Groton, Connecticut 06340

Lt. Col. Austin W. Kibler
Director, Human Resources Office
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, Virginia 22209

U.S. Air Force Office of Scientific
Research
Life Sciences Directorate, NL
1400 Wilson Blvd.
Arlington, Virginia 22209

Dr. J. M. Cristensen
Chief, Human Engineering Division
Aerospace Medical Research Laboratory
Wright-Patterson AFB, OH 45433

Dr. J. E. Uhlaner
Dir., U.S. Army Research Institute
for the Social & Behavioral Sciences
1300 Wilson Blvd.
Arlington, Virginia 22209

Chief of Research and Development
Human Factors Branch
Behavioral Science Division
Department of the Army
ATTN: Mr. J. Barber
Washington, D. C. 20310

Dr. Joseph Zeidner
Dir., Organization and Systems
Research Laboratory
U.S. Army Research Institute for
the Behavioral & Social Sciences
1300 Wilson Blvd.
Arlington, Virginia 22209

Dr. Stanley Deutsch
Chief, Man-Systems Integration
OART, Hqs., NASA
600 Independence Avenue
Washington, D. C. 20546

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Virginia 22202

Dr. Edgar M. Johnson
Organizations & Systems Research Lab.
U.S. Army Research Institute for the
Behavioral and Social Sciences
1300 Wilson Blvd.
Arlington, Virginia 22209

Dr. James Parker
BioTechnology, Inc.
3027 Rosemary Lane
Falls Church, Virginia 22042

Dr. Edwin A. Fleishman
Foxhall Square
3301 New Mexico Avenue, N.W.
Washington, D. C. 20016

American Institutes for Research Library
135 N. Bellefield Avenue
Pittsburgh, Pennsylvania 15213

Psychological Abstracts
American Psychological Association
1200 17th Street, N.W.
Washington, D. C. 20036

Dr. A. I. Siegal
Applied Psychological Services
404 E. Lancaster Street
Wayne, Pennsylvania 19087

Dr. Joseph Wulfeck
Dunlap and Associates, Inc.
115 South Oak Street
Inglewood, California 90301

Dr. Robert R. Mackie
Human Factors Research, Inc.
Santa Barbara Research Park
6780 Cortona Drive
Goleta, California 93017

Mr. Wes Woodson
Man Factors, Inc.
4433 Convoy Street, Suite D
San Diego, California 92111

Dr. C. H. Baker
Director, Human Factors Wing
Defense & Civil Institute of
Environmental Medicine
P. O. Box 2000
Downsville, Toronto, Ontario
Canada

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, N.W.
Washington, D. C. 20036

Dr. Bruce M. Ross
Department of Psychology
Catholic University
Washington, D. C. 20017

Mr. Harry Chipman
WR Systems, Inc.
2531 S. Jefferson Davis Highway
Arlington, Virginia 22202

Dr. David Meister
U.S. Army Research Institute
1300 Wilson Blvd.
Arlington, Virginia 22209

Mr. George Graine
Naval Ship Systems Command
(SHIPS 047C12)
Department of the Navy
Washington, D. C. 20362

Lt. Col. John F. Ahlborn
Headquarters, AFSC-DLSE
Andrews AFB
Washington, D. C. 20334

Dr. H. H. Wolff
Technical Director (Code N-2)
Naval Training Equipment Center
Orlando, Florida 32813

Dr. Donald A. Topmiller
Chief, Systems Effect. Branch
Human Engineering Division, USAF
Wright Patterson AFB, Ohio 45433

Col. Robert O. Viterna
DA/OCRD
Headquarters, Dept. of the Army
DARD-ARS-B
Washington, C. C. 20310

Dr. Anthony Debons
IDIS
University of Pittsburgh
135 N. Bellefield Avenue
Pittsburgh, Pennsylvania 15260

Dr. Alfred F. Smode
Training Analysis & Evaluation Group
Code N-OOT
Orlando, Florida 32813