DOCUMENT RESUME

ED 112 871 IR 002 581

AUTHOR	Mann, William C.; And Others
TITLE	Observation Methods for Human Dialogue.
INSTITUTION	University of Southern California, Marina del Rey. Information Sciences Inst.
SPONS AGENCY	Advanced Research Projects Agency (DOD), Washington, D.C.
REPORT NO	ISI-RR-75-33
PUB DATE	30 Jun 75
NOTE	92p.
EDRS PRICE	MF-$0.76 HC-$4.43 Plus Postage
DESCRIPTORS	Artificial Intelligence; *Communication (Thought Transfer); Communications; Computational Linguistics; *Computer Science; Data Collection; Human Development; Humanization; Information Theory; *Man Machine Systems; *Observation; Research Methodology
IDENTIFIERS	Human Dialogue; Natural Language

ABSTRACT

Progress is described on a new approach to improve man-machine communication in order to significantly expand and diversify the capabilities of the computer interfaces that people use. The objective is to design computer processes that can assimilate particular aspects of dialogue between people, then transfer the processes into man-machine communication. Nine aspects of human ability to communicate are selected and studied in detail, according to new methods of data collection described in the report. Explicit observational instructions also are given. (Author/SK)

ARPA ORDER NO. 2930
NR 154-374

ISI/RR-75-33
June 1975

William C. Mann
James A. Moore
James A. Levin
James H. Carlisle

Observation Methods for Human Dialogue

INFORMATION SCIENCES INSTITUTE
UNIVERSITY OF SOUTHERN CALIFORNIA
4676 Admiralty Way/Marina del Rey/California 90291
(213) 822-1511

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.
The views and conclusions contained in this document are those of the author(s) and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Office of Naval Research, the Defense Advanced Research Projects Agency, or the U.S. Government.
This document is approved for public release and sale; distribution is unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ISI/RR-75-33
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): OBSERVATION METHODS FOR HUMAN DIALOGUE

5. TYPE OF REPORT & PERIOD COVERED: Technical - January 1, 1975 through June 30, 1975

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): William C. Mann, James A. Moore, James A. Levin, James H. Carlisle

8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0710

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90291

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
61153N RR042-06-01
RR042-06 NR154-374

11. CONTROLLING OFFICE NAME AND ADDRESS:
Human Resources Research Office
Advanced Research Projects Agency
1400 Wilson Blvd., Arlington, VA 22209

12. REPORT DATE: June 30, 1975

13. NUMBER OF PAGES: 90

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
Personnel and Training Research Programs
Office of Naval Research - Code 458
800 N. Quincy Street
Arlington, VA 22217

15. SECURITY CLASS. (of this report): Unclassified

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):

18. SUPPLEMENTARY NOTES:

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
dialogue, man-machine communication, computer, process model, natural language, data collection, human communication, artificial intelligence, linguistics, cognitive psychology, computer science, group behavior, coding methods, reference, topic, question, request, comprehension, understanding, correction, paraphrase.

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

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

This report describes progress on a new approach for improving man-machine communication. The goal of the work is to significantly expand and diversify the capabilities of the computer interfaces that people use. The approach is first to design computer processes that can assimilate particular aspects of dialogue between people, then to transfer these processes into man-machine communication.

The approach requires that particular aspects of the human ability to communicate be selected and studied in detail. This report describes new methods of data collection developed to meet this need and tells how they will be used.

The report focuses on nine phenomena of human dialogue which have been selected from approximately 23 phenomena proposed and explored. For most of the nine, explicit observational instructions are given as well.

OBSERVATION METHODS FOR HUMAN DIALOGUE

by

William C. Mann
James A. Moore
James A. Levin
James H. Carlisle

ISI/RR-75-33

June 1975

USC/Information Sciences Institute
4676 Admiralty Way
Marina Del Rey, California 90291

The research reported herein is supported by the Personnel and Training Research Programs of the Office of Naval Research (Code 458), Arlington VA 22217, Contract Authority Identification Number NR 154-374, Contract N00014-75-C-0710, under ARPA Order Number 2930 from the Human Resources Research Office of the Advanced Research Projects Agency. The report is approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government. Opinions expressed herein are those of the authors and do not necessarily reflect policy of any agency of the United States Government.

CONTENTS

ABSTRACT

This report describes progress on a new approach for improving man-machine communication. The goal of the work is to significantly expand and diversify the capabilities of the computer interfaces that people use. The approach is first to design computer processes that can *assimilate* particular aspects of dialogue between people, then to transfer these processes into man-machine communication.

The approach requires that particular aspects of the human ability to communicate be selected and studied in detail. This report describes new methods of data collection developed to meet this need and tells how they will be used.

The report focuses on nine phenomena of human dialogue which have been selected from approximately 23 phenomena proposed and explored. For most of the nine, explicit observational instructions are given as well.

1. OVERVIEW OF THE REPORT

This report is a progress report rather than a report of completed results. While enough has been accomplished to be interesting, it is all tentative and subject to change; therefore we simply exhibit it.

Most of the report is a description of methods which a person, called an Observer, can use to identify and describe nine kinds of phenomena which arise in human dialogue. The phenomena are:

Repeated Reference
Requests
Expression of Comprehension
Similar Expressions - two kinds
Topic Structure
Incremental and Prerequisite Content
Correction Actions

The methods are used in building computer programs that can interpret dialogue. We expect that the subprocesses of these programs will be able to improve man-machine communication when they are implanted in task-oriented systems.

More context than we can supply here is needed for an accurate interpretation of the significance of the observation methods. The context-setting sections do not tell the whole story, but we expect that most readers only want a brief overview. The report does not contain a detailed presentation of the relationship of this work to preceding and concurrent research. This has been done elsewhere (Mann 1974) and will be updated in forthcoming publications and reports.

The work reported here is not embedded in a traditional, well-worn approach to communication. It is a new approach to the study of person-to-person communication, with particular attention to improving man-machine communication as a consequence. The report deals almost exclusively with data-definition and data-collection portions of the methodology, which are but a small part of the whole. The approach draws heavily on computer science, linguistics, and psychology, and on other disciplines to a lesser extent.

In yet another way this report is not representative the whole. It suggests a particular broad scope of attention for which we are now prepared to do certain kinds of data development. But we know that the whole scope cannot be approached at once; it must be done gradually, and the selectivity of that movement is not at all apparent in the report.

Finally, some cautions are in order about the Observation Instructions in the report. They are not really ready for direct use by anyone other than the authors. In developing them, primary attention was paid to what kinds of things to annotate rather than to the notational forms or the presentation of ideas to the Observer. The major reason for this is that we expect the substance of the instructions to change significantly under the pressure of the next few months of use. Any significant cosmetic or presentational cleanup at present would be premature, as would any attempt to determine their reliability.

As a progress report, it is a contribution to empirical research on human symbolic communication, and thereby to man-machine communication.

During the past two years, there has been a very substantial increase in the number of on-line, interactive users of computers in the military. More important, many of the long-term military aspirations for command-control and for administrative and support functions call for computer-based, on-line, interactive systems. Thus this research is meeting a growing need for effective, user-oriented, man-machine interfaces directly and effectively supporting military requirements.

2. THE RESEARCH ENVIRONMENT OF THE OBSERVER'S METHODS

In this section we will relate the Observer's task to our general methodology. Each Observer (working with a single dialogue) carries out one instance of the basic experiment, which in turn will lead to a model for that particular dialogue. This basic experiment is to be repeated many times with different Observers and dialogues. Those dialogue-comprehension processes which have repeatedly shown themselves valuable throughout these experiments will then constitute the overall results; they are the primary candidates for the final step of embedding in task-oriented system applications.

THE BASIC EXPERIMENT

The basic experiment consists of four steps, as illustrated in Figure 2-1: 1) acquire transcript of dialogue for study, 2) gather Observer's commentaries, 3) construct a process model to account for these observations, and 4) compare the actions of the processes in the model with the Observer's judgments and Observer's behaviors.

Dialogue Transcripts

We have chosen to deal with dialogues only; we do not plan to cover cases of multi-person conversation. The dialogue must be in a machine-readable representation. We will use a conventionally typed transcript of the text of the dialogue, perhaps with

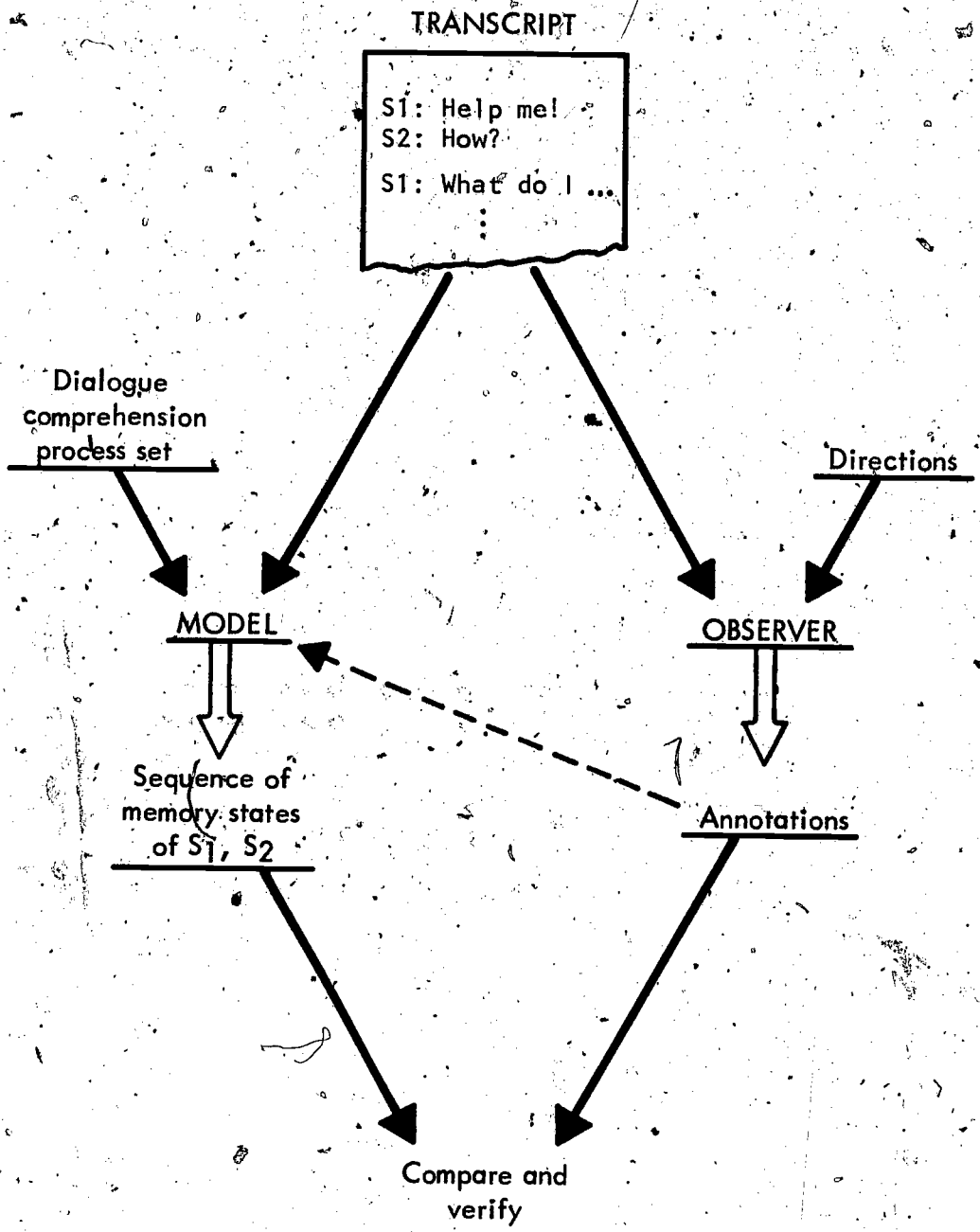


Figure 2-1 - Basic Dialogue Case Experiment

some timing information. We do not now envision any attempt to include and use such subtleties as facial expression, gestures, and intonation. In order to assure ourselves that the transcripts are not systematically excluding aspects which were significant in the original conversation, we will examine only those dialogues which were originally conducted over separated media (e.g. teletype, voice radio, etc.) where the participants successfully communicated, despite the lack of these other, rich sources of information.

We currently have several hundred transcripts of dialogues from a wide variety of sources. These include:

1) Help-seeking dialogues from the TENEX time-sharing system.
2) Astronaut-and-Ground-Control dialogues from the Apollo-13 mission.
3) Tutoring transcripts from various sources.
4) Transcripts of radio talk shows.

We have been using some of these to pretest the directions to the Observers. Some of these transcripts fail our criteria in one or more ways, often because of occasional intrusions by a third party. We are also developing a facility to gather transcripts ourselves, directly from a terminal-to-terminal dialogue.

Since there are certain phenomena we definitely do not plan to model, and since we are committed to giving precisely the same transcript to the model as to the Observer, we will perform a certain amount of "cleaning up" of the transcript to get rid of the features which we know the model will not be able to deal with. In particular, we are not modeling the participants' ability to deal with mispelings, abbrev'ns, and wordsruntogether, the various ways in which a typist (especially on-line to a computer) will indicate local correx\Xctin\Nons, etc. These and similar kinds of "noise" will be

corrected before either the model or the Observer see them. What will *not* be cleaned up is the anomalous grammar, inappropriate choice of words, saying things which clearly (to us) weren't meant, and the assorted fumblings with words which happen when the speaker can't find the right way to express himself.

Additional details of how a transcript is prepared for the Observer are given in the Appendix, with examples.

Observer Commentaries

Having chosen a particular dialogue for study, we next give it to an Observer, chosen from outside our research team, to obtain his commentary on the dialogue. This activity is described in detail in sections 5 through 12 of the report. (Thus the report covers methods for a part of the experiment cycle rather than the entire cycle.)

Dialogue Process Models

With both the transcript and the Observer's annotations to guide us, the next step is to build a set of processes which will maintain a pair of simulated memories (of the information states of the two participants) based on the transcript as input. Although we will feel free to reuse processes from the models in our previous experiments when convenient, our intentions are that each of these models be a one-of-a-kind program, responsible only for the one dialogue and one Observer on which it was based. This view contrasts with the more conventional approach of building a single system which, with each iteration of the experimental cycle, would be expected to adequately model the new dialogue/commentary pair *as well as* all the preceding ones.

Obviously, certain supporting processes will be needed, regardless of the choice of transcript. Among these might be processes for:

1. Natural language parsing
2. Semantic memory management
3. Inference
4. Discrepancy detection and resolution
5. Hypothesis generation and testing
6. Evidence evaluation
7. Awareness of time passage in dialogue
8. Attention focus
9. Selective forgetting
10. Time and space resource allocation
11. Extensive trace and debugging facilities

These prospective supporting processes share common technical histories. Each has precedents in existing computer programs, although the combination does not. Each one is subject to simplification (relative to past instances) because our models are case models rather than general systems. For each, there exist adequate technology and personal skills for building the necessary experimental programs. Nearly all of this expertise comes from artificial intelligence, cognitive psychology, and computational linguistics.

In any model of a single dialogue, there is an issue of what parts of the model are nonsignificant because they are ad hoc. This problem is addressed explicitly at the multiple experiment level.

Comparison of Process Behavior with Observer Commentary

Once we have constructed the model we will conduct an extensive comparison of the behavior of the model with the Observer's commentary. We expect to identify three sets of results from each experiment:

1. Program states and actions corresponding to the Observer's commentary.
2. Phenomena recorded in the Observer's comments for which there are no corresponding states in the program.
3. Program states which appear to contradict the Observer's commentary.

In addition, for those aspects of the model which bear a positive relationship to the Observer's commentary, we will estimate the (actual or potential) generality of the methods involved.

For behavior noted by the Observer but contradicted (or ignored) by the model, we will analyze why it was difficult or undesirable to achieve accuracy in the model along these dimensions.

MULTI-EXPERIMENT PROCESS EVALUATION

As we repeat the above experiment on new dialogues with new Observers, we expect to accumulate a collection of processes which are reusable as is, or with minor generalization. Those processes which continue to prove useful over a collection of experiments are the ones which we will consider candidates for export. (See Figure 2-2)

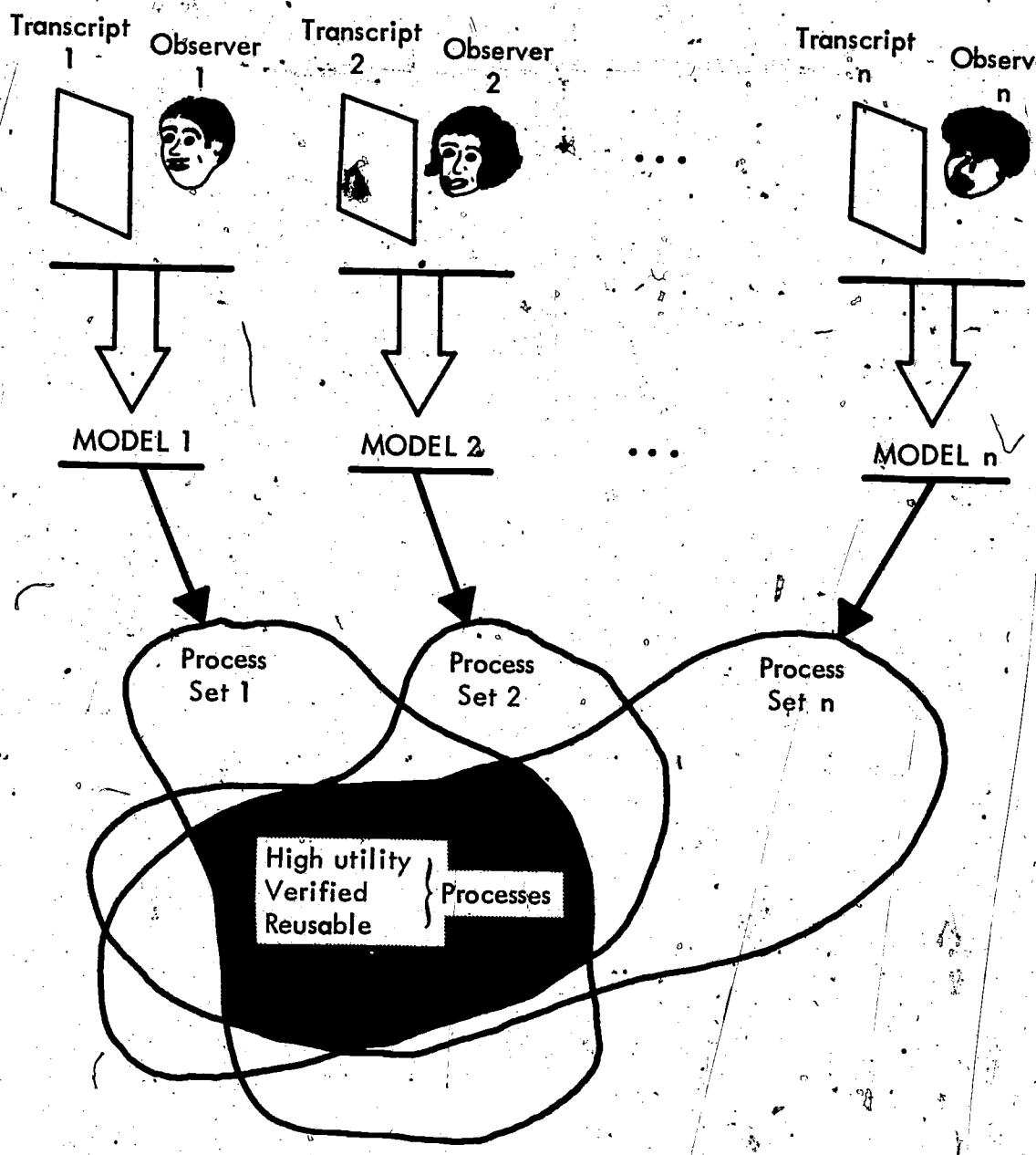


Figure 2-2 - Process Evaluation in Multiple Experiments

This is the answer to the problem of adhocness. Processes are not identified in this way unless they continue to work over a range of dialogues. It is a conservative answer, in that some processes that in fact are generally applicable may not be exercised enough to be recognized in this step. The range of cases over which a process has been tested is always explicit, so that the evidence for its generality is clear.

One of the advantages of this multistage approach is that we expect certain processes to prove themselves useful early and thus to provide us with exportable results well before the end of the project.

DISSEMINATION OF DISCOVERED ALGORITHMS

Dissemination of our results is uniquely important for this project. By the very nature of our goals, our work is broadly interdisciplinary, with implications for computer science, psychology, linguistics, and other disciplines. The technical dissemination will include the novel features of the methodology and the novel result forms as well as specific processes.

Two kinds of dissemination of results are planned: one conventional and one specific to this method. The first kind of dissemination will occur through the usual scientific channels: site visits, scientific and popular publication, conference attendance and presentations, as well as exchange of programs and, perhaps, personnel. When we have established a collection of repeatedly effective processes, we plan to disseminate these results by actually retrofitting the processes to already existing man-machine interfaces. Of course, they can be designed into new systems as well.

3. UNDERSTANDING THE OBSERVER'S TASK

Two kinds of input information are available to the Observer as indicated in Figure 3-1. He receives a dialogue transcript and a set of instructions on how to annotate it. He will also receive training and practice in the use of these instructions.

The instructions include a general orientation to the experimental goals, the uses of the Observer's work, and the ground rules for making judgments; they also include directions for several specific annotation tasks to be performed on copies of the transcript. These tasks are performable independent of each other except for a few specific sequence requirements. These annotations are the only output produced by the Observer.

The task is a blend of specified, definite steps and personal, subjective judgments. For example, we ask the Observer to tell where each participant has expressed comprehension of the other's remarks, but we do not tell him *how* he is to know this.

The subjective parts of the Observer's work are essential because they engage his communication abilities and methods. Certain methods are used in both his own performance of his task and in the dialogue participants' performance. It is these methods, involving receptive acts in communication, that will be represented in our models and their comparative evaluation.

The skills used by the Observer are primarily those of an ordinary native-communicator in the language of the dialogue being observed (which always

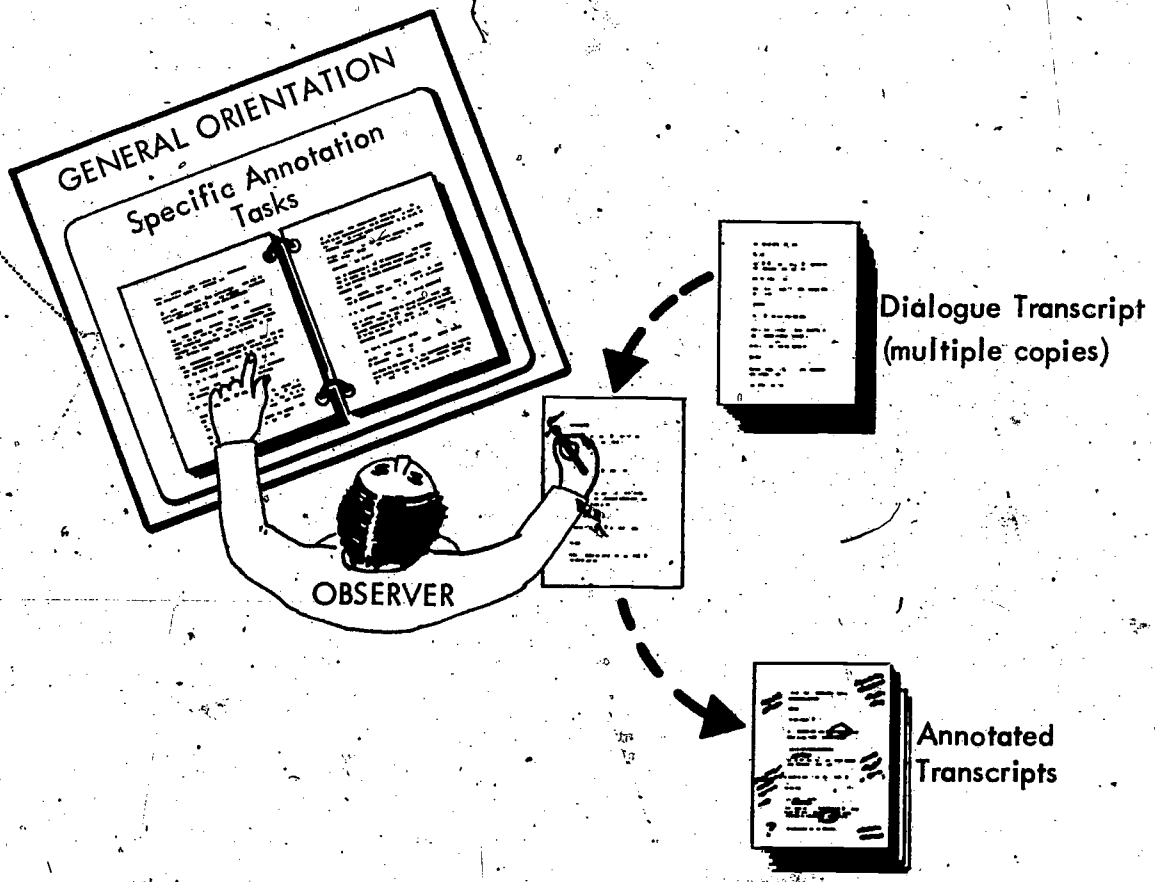


Figure 3-1 - The Observer's Task

happens to be English). One Observer is used for all of the categories; the annotation task is not split across Observers, and it does not include pooled final judgments. This assures that the view of a dialogue that gets built into a model can be as coherent as the individual Observer was. (It avoids models resembling committee reports.) The Observer makes judgments that fo him are obvious, clear case judgments, so that the adequacy of the evidence in the transcript is thereby assured. His work does not require sophisticated technical understanding of language, and it does not include computer-related notation in any way. (In strict experimentation the Observer, the model builders, and the dialogue participants are always separate.) We expect that his judgments are representative of those made by the great majority of native-communicators in the course of their actual communication.

The Observer will be supplied with information on the situation, roles, jargon, and medium of the dialogue. The general rule is that his access to each turn of the dialogue should be essentially as good as that of the participants.

Our policy is not to deceive the Observer in any way. Beside our preference for this approach on moral grounds, it has a number of technical benefits as well. We believe that the Observer's performance will be more stable if he is not attempting to discover the "real" experiment. We also expect to disseminate the technology to interested volunteer Observers, which requires providing them with the best possible access to it. Also specialists who may serve as Observers from time to time. Finally, effective deceit is difficult to achieve, and experiments which rely on it are inherently suspect.

4. PHENOMENA FOR OBSERVATION

THE OBSERVATIONS

We have developed and pretested instructions for our Observers to comment on a number of interesting dialogue phenomena. These include recognition of references both to previously expressed concepts and to previous text. We also have instructions for certain common dialogue forms, such as questions and commands, expression of comprehension (and confusion), and the correction of errors and misunderstandings. We have a number of ways to capture aspects of the content conveyed in a dialogue and the knowledge relied upon to comprehend this content. This includes ways of generating and judging paraphrases of dialogue utterances that permit a detailed examination of the role of context in dialogue. In addition, we have developed instructions for Observers to annotate the topic structure in dialogue as introduced and terminated by each participant.

CRITERIA FOR SELECTING PHENOMENA

The various phenomena described briefly above (and in detail below) by no means encompass all aspects of dialogue. There are a number of criteria that we have informally used to select phenomena to observe.

1) Importance:

First of all, we have considered only phenomena that we felt were an important component in dialogue. Each set of instructions sheds light on an aspect of dialogue that we feel will be central to an interesting model of two-party communication.

We have required that phenomena be central, in the sense that the communication would break down or be significantly changed in character if the phenomenon in question were eliminated. Phenomena related to a particular dialogue source or context were not chosen, and phenomena which appear substantially unaltered in monologues were not chosen.

2) Clarity:

In the course of writing these instructions, we developed and modified our concepts of the phenomena, since we were forced to specify in some detail what we wanted to observe. The original five categories of phenomena to observe specified in the original description of our work (Mann, 1974) were developed into the present nine categories through this classification process. The instructions in this report have typically been through three or four major revisions to improve clarity. Some of the phenomena that we consider important will not be included in immediately forthcoming experiments because we have not so far been able to write clear instructions for their observation.

3) Reliability:

Finally, the clear notions of what to observe had to be translated into instructions that produced reasonably consistent observations across Observers in our pretests. Our informal feeling of the present level of reliability can be expressed as follows: One could define appropriate measures of agreement on these annotations, fairly and reasonably, with which the authors would score above 90 percent on most of these categories, working on the kind of dialogues involved in the pretest.

The Development of Observation Categories

In the process of developing the nine categories presented in this report, we examined fourteen other observation categories in detail.

Three of these other categories were dropped because they failed one of the criteria described above. This was usually because the concepts could not be operationalized into coding rules for Observers.

In many other cases, we were able to shift from initially unusable categories to others that bore on the same phenomena in a way which met our constraints. Seven categories developed into other categories in this way.

Finally, four categories of observation are still in a state of development. These categories are promising, but have not yet developed to the stage where they can be presented here. These include observations of ceremonies that occur in dialogues, of comments about the dialogue itself, of the speaker's functions (or purposes) in generating his utterances, and of defining events in which the meanings of new terms or concepts are explicitly given.

So the set of observation categories described in this report are those members of a larger set that survived the criteria above. We would expect to add to this set in the future after experience with modeling and observation brings new aspects of dialogue to light.

POST-PROCESSING

As described previously, the commentary of the Observers will be used to test the dialogue models we build. This makes it necessary to compare the commentary to the behavior of the model. In some cases, this comparison is straightforward. For example, when the Observer annotates that two sets of words refer to the same concept, it will

be easy to tell whether the model also creates this correspondence. In other cases, some post-processing of the Observer's annotation will be required before we can make a comparison. For example, we have derived a simple algorithm for specifying the areas in which a topic is shared by both participants, given the annotation of topic initiation and termination for each participant separately. We expect to develop further post-processing algorithms to enable us to test parts of the model for which we have no direct observation methods.

FORMAT OF THE INDIVIDUAL INSTRUCTIONS SECTIONS

The next sections will discuss the instructions developed so far. For each set, we will present the phenomena to be observed, the instructions themselves, and an example of applying the instructions to a real dialogue.

The dialogue used for each example is an interaction between a computer user and the computer operator, conducted remotely via computer terminals. Thus the transcript, which is given in full in the Appendix, captures almost all the interaction (only the timing specifications are lacking).

5. ORIENTATION FOR THE OBSERVER

The Observers' directions are presented in several sections, one for each of the classes of phenomena we are interested in having annotated. There is a small set of instructions which precede and are common to each of these sections. For the sake of brevity these are presented here, once, rather than at the head of each separate set.

OBSERVER'S GENERAL INSTRUCTIONS

We are interested in your commentary on those features and phenomena which you regard as clear and obvious. Do not worry about making subtle or questionable judgments. For those instructions of the form: "Identify those parts of the dialogue having the property ... ", if you have any doubts whether a certain segment meets the criteria set forth in the directions, just ignore it. Some instructions will have the form: "For a piece of dialogue already identified, classify it into one of the following categories: ... ". These sets of categories will always have one labelled: "other/unknown/unsure"; use this classification freely if you aren't sure that one of the other classifications is appropriate.

Some of the transcripts we are working with have been typed by a secretary, listening to a recorded dialogue; others have been typed by the participants themselves as they were conducting the dialogue via terminals. You will be told which of these cases applies to the dialogue you are annotating. This is only important to you in one regard: punctuation. If the transcript was prepared by anyone other than the actual participant in the dialogue, then the punctuation you see is to be taken as only a "good guess," possibly in error. On the other hand, if the transcript was created by the participants' typing, then the punctuation can be taken to be what the participant actually expressed, with no editorial interpretation involved.

Although a few of the directions ask you to differentiate between the knowledge of the two participants, most do not, since it is assumed that both participants understood the conversation in substantially the same way. In the case where the instructions clearly presume that the participants share an understanding of what is being communicated and you perceive that any of your annotations reflect only the understanding of *one* of the participants, indicate this by adding the name (letter) of that participant to the annotation.

A number of the directions call for you to indicate a certain segment of the dialogue having some property. You may find that the segment has sharp boundaries, or they may be somewhat indefinite. For all these directions which call for you to mark these segment boundaries (with what we will call "segment markers") the following general instructions apply:

Put a "[" at the point where a segment abruptly starts, and put a "]" at the point where a segment suddenly ends.

Graphically:

```
                            |
                            |
      A: text text text  [  text text text
                            |
         <<<<<<<<<<<<<<<<<<<|>>>>>>>>>>>>>>>
         (definitely        | (definitely
          NOT part          |   IS part
          of segment)       |   of segment)
                            |
```

After you finish marking these abrupt boundaries, go back to the beginning of the dialogue and determine where gradual starts and ends occur. Mark a "(" at the first point that a segment is clearly under way, and a "(=" at the latest point (prior to the "(") that clearly is not included in the segment. Similarly for gradual segment ends, mark a ")" at the last point that the segment is clearly under way, and a "=)" at the first point (following the ")") that the segment clearly does not include.

Graphically:

```
                               |                      |
                               |                      |
B: text text text text  )  text text text text  =)  text text text
   <<<<<<<<<<<<<<<<<<<<<|                           |>>>>>>>>>>>>>>>>>>>>
   (definitely IS part  |  (unclear whether         |  (definitely not
    of segment)         |   or not is part          |   part of segment)
                        |   of segment)             |
                        |                           |
```

Notice that each use of "[" is equivalent to a simultaneous occurrence of "(=" and "(".

Several of the instructions will call for you to create labels for segments of the dialogue; these labels will consist of one or more letters followed by a number. This choice of numbers is completely arbitrary -- you need not be concerned with keeping them either consecutive or in order (or corresponding to the numbers used in compliance with other sets of directions).

Whenever you feel that the actual annotations are not capturing the phenomena to which they are addressed, please note this and add your own comments at any point.

6. *REPEATED REFERENCE AND TEXT REFERENCE*

One of the most common phenomena in language is that of reference, the use of a sequence of words by a speaker to refer to an object or concept. Taken as a whole, the concept of reference is extremely complicated, encompassing as it does almost the entire range of human conceptual and experiential abiiities. Not surprisingly, we do not aspire to model the full range of reference phenomena in natural language.

We have made two major restrictions on the phenomena we are investigating: First, we are only interested in those cases for which the same object or concept is referred to by a participant in more than one place in the dialogue. Second, we do not attempt to pin down exactly what is being referred to--rather, we simply want to determine which sequences of words have been used to refer to the same thing (whatever it may be).

Within these boundaries, we investigate a few of different flavors of reference:

1. Two regions of dialogue refer to the same thing.
2. One region refers to a set of things and a second region refers to a single thing which is a member of that set.
3. One region refers to a set of things, a subset of which is referred to by the second region.

We are also investigating a form of reference with a somewhat different thrust: a so-called "Text Reference," made to a string of words in the preceding dialogue itself (and not to the referent of that preceding string of words!).

OBSERVER'S DIRECTIONS FOR REPEATED REFERENCE AND TEXT REFERENCE

Repeated Reference

A ***Repeated Reference*** is said to occur whenever two phrases in a dialogue are used to refer to the same thing (object, person, activity, concept, ...). Some examples of such ***Repeated References*** are the following:

1. Repeated Reference to an object
   A: I really like your wrist watch.
   B: It only cost $50.00.
2. Repeated Reference to a person:
   A: I spoke to Max's sister yesterday.
   B: Did you know that she's going to Europe next month?
3. Repeated Reference to an activity
   A: Sky diving is a great sport.
   B: Yes, and it's not as dangerous as some people think.
4. Repeated Reference to a concept
   A: Jim told me that jobs are very hard to find now.
   B: Yes, that's true. It certainly is sad...

The two phrases may be identical in wording, but in general are not. A ***Repeated Reference*** may be effected by as little as a single word, or as much as a complete utterance. A phrase may participate in many ***Repeated References*** or in none.

Whenever in your judgment two phrases are obviously used to refer to the same thing 1) underline each phrase, 2) assign the first a unique number, 3) write this number in either margin, opposite each of the two phrases, 4) connect each phrase to the number you just wrote with a line.

U: Is [NAME5] still around? (1)

O: No he isn't. (1)

If the same thing is referred to by means of more than two phrases, use the number assigned to the first such phrase in annotating each subsequent Repeated Reference.

U: Can you recover those files for me?

(7)

O: OK hold on just a min. and I'll try to find them

(7)

...

O: OK yes I have found the files you were concerned about

(7)

You may discover that one phrase is used to refer to a single element of the set or collection of things referred to elsewhere by means of a second phrase. Whenever this happens, annotate these as ***Repeated References*** with the addition of an "E" or "S" (respectively) on the line from the Reference to an element or to a set. For example:

O: I am really not sure, however I know that that would

E(6)

not be responsibility we would take on ourselves ...

(6)S

Sometimes one may find that one phrase is used to refer to a set and a second phrase is used to refer, not to an element, but to a subset of the set referred to by means of the first phrase. In such cases of ***Repeated References***, mark the Reference to the set with an "S" (as requested above), and mark the Reference to the subset with "SS". For example:

U: Are the files that are on disk archived every week?

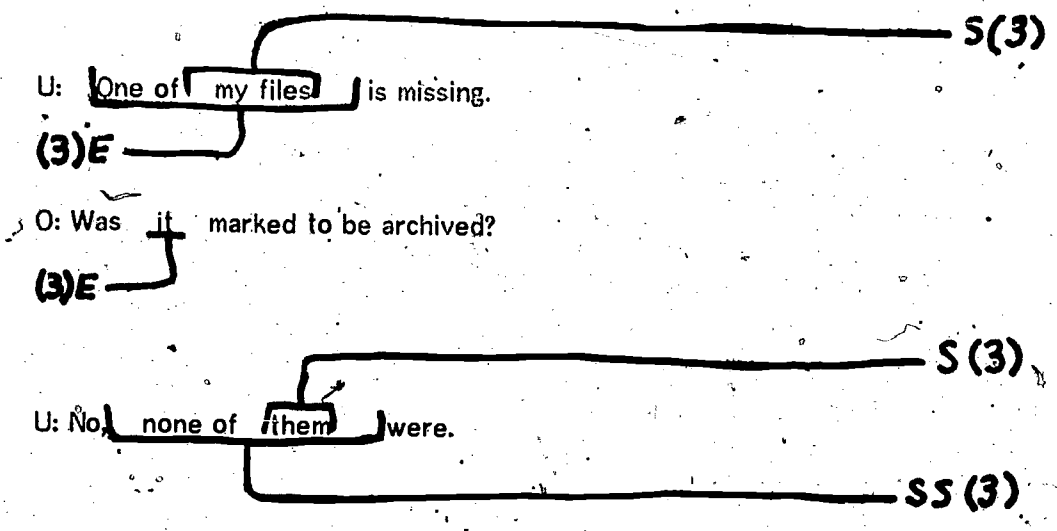
(32)S

O: Yes.

U: I had some files in one day and the next they were gone.

SS(32)

In the case of one Reference within another, underline and bracket the larger one; overline and bracket the smaller one. For example:



Referring expressions are a very common phenomenon in language--so much so that some are easily overlooked To help you be sure that you are considering all possible ***Repeated References,*** pay particular attention to phrases beginning with

1. Personal Pronouns
   (I you he she it we they me him her us them) N.B.: these will be one-word phrases.

2. Quantifiers
   (a an the that each all some any every one ...)

3. Pronominal Possessives
   (my your his her our its their mine yours hers ours theirs) N.B.: these phrases may have two ***Repeated References***: one for the possessor and one for the object possessed.

4. Wh-words
   (who what when where why which how ...)

These words are intended to serve as clues to most of the potential ***Repeated References*** but are not intended to be comprehensive. ANY two sets of words which you feel refer to the same thing are to be treated as a ***Repeated Reference.***

Since they are so common in dialogue, we are making a special case of first and second person pronouns. All occurrences of first and second person

SINGULAR pronouns (I me my mine you your yours), when used to refer to self and partner, respectively, are simply to be underlined, with no further annotation necessary. Should you find an occurrence of a plural pronoun in the first or second person, select one convenient occurrence of a singular pronoun used by the same speaker, in the same person. Annotate these two (the occurrences of the singular and plural pronouns) as per the instructions for general "element" and "set" References, above. Note that first and second person singular possessive pronouns may still require annotation as a Reference to the thing possessed, but not as a Reference to the possessor.

You should be careful to distinguish (where possible) between "you" referring to the speaker's partner and "you" referring to an indefinite "other" (e.g., as in "You can't fight City Hall"). When you come across a clear instance of an impersonal or general "you", do not underline it; rather, indicate the occurrence of an impersonal "you" by circling it in the dialogue. If you are not sure whether or not a particular occurrence of "you" is impersonal, simply do not annotate it at all - neither underline it nor circle it.

Text Reference

A ***Text Reference*** occurs whenever reference is made to previously occurring words within the transcript. For example, in the sentence

> Go 3 blocks and turn North;
> By North I mean towards the mountains.

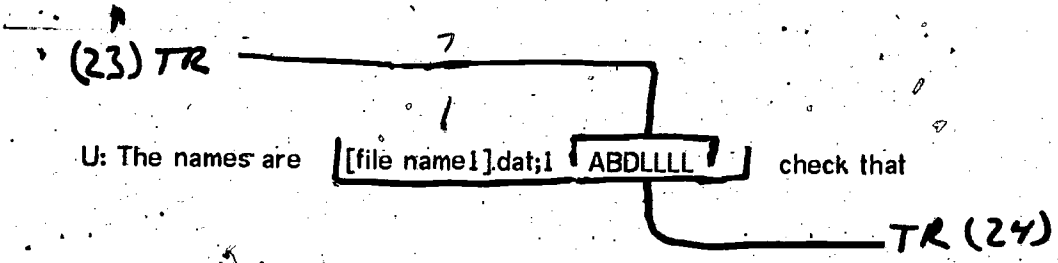
the second use of "North" is a ***Text Reference*** to the first. We call this a

Text Reference because it refers to the previous use of the word itself, rather than to its meaning.

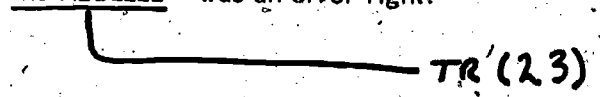
Many of what we are calling ***Text References*** would be indicated, in a formal publication, by quotation marks. However, the transcripts typically will not include these quotation marks. Note that a ***Text References*** need not repeat the words to which reference is being made. For example:

In your last three sentences, you failed to answer me.

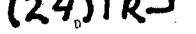
Mark ***Text References*** in the same manner as ***Repeated References***, and distinguish them with a "TR" next to the number. For example:



[file name1].dat;1,2 [file name2].f4;1 and [file namr3].f4;1



U: Right that should have been [file name].dat; 1 and 2.



Sometimes there will be a ***Text Reference*** to all or one of a number of separate pieces of text, each involving the same phrase or word. You need annotate only the most recent, prior occurrence of the text which seems to have the same meaning. For example:

E: An interpreter is always faster than a compiler.
(17)

A: Why?

E: An interpreter starts producing output immediately.
(17) TR (18)

A: What do you mean by an interpreter ?
(17) TR (18)

Be sure you examine all occurrences of:

- SAY
- MEAN
- MENTION
- TALK ABOUT
- DEFINE

and similar phrases (and their other tenses) to see whether they signal the presence of *Text References*. As before, this list is is only suggestive and is not presented as being comprehensive.

Figure 6-1 represents a portion of a dialogue fully annotated for *Repeated Reference*.

15 O: YES

16 U: IN ONE OF MY DIRECTORIES I HAD SOME FILES IN ONE DAY AND THE NEXT THEY WERE
17 GONE AND THEY WERE NOT ARCHIVED

18 O: WELL, I AM REALLY NOT THAT SURE OF HOW THE ARCHIVEING OF
19 THE FILES ARE DONE SO I WOULD SUGGEST THAT YOU SEND A MESSAGE TO
20 [name2] IN REGARDS TO THIS PROBLEM. THE ARCHIVEING IS DONE
21 DURING THE SWING SHIFT SO IN ORDER THAT YOU DO DON'T NOT GET
22 MISLED BY SOMETHING I AM NOT CERTAIN ON I WOULD RATHER
23 YOU DISCUSS THIS WITH [name2]. HE IS OUT OF HIS OFFICE AT THIS
24 TIME BUT YOU COULD SEND HIM A MESSAGE AND I AM SURE HE WILL GIVE
25 YOU ALL THE INFORMATION YOU WOULD LIKE. [operator's name]

26 U: OK [operator's name]. CAN YOU PEOPLE GO INTO ANY DIRECTORY AND DELETE FILES WHEN YOU
27 NOTICE THAT THEY ARE NOT BEING USED. [user's name]

28 O: I AM REALLY NOT SURE, HOWEVER I KNOW THAT THAT WOULD NOT BE
29 RESPONSIBILITY WE WOULD TAKE UPON OURSELVES TO JUDGE WHETHER OR
30 NOT YOU WANT YOUR FILES OR NOT.

31 [outside interruption here]

32 O: [user's name], ARE YOU THERE?

Figure 6-1a - Dialogue Annotated for Repeated Reference - Part a

33 LINK FROM [user's name], JOB 21, TTY 22

34 U: SORRY [operator's name]

35 WE GOT DISCONNECTED....CAN YOU RECOVER THOSE FILES FOR ME..

36 AS FAR AS I KNOW THEY WERE IN THE DIRECTORY ON THE 16TH...THE NAMES

37 ARE ...[file name1].DAT;1ABDLLLL CHECK THAT ...[file name1].DAT;1,2

38 [file name2].F4;1....AND [file name3].F4;1.....[user's name]

39 O: OK HOLD ON JUST A MINUTE AND I WILL TRY TO FIND THEM

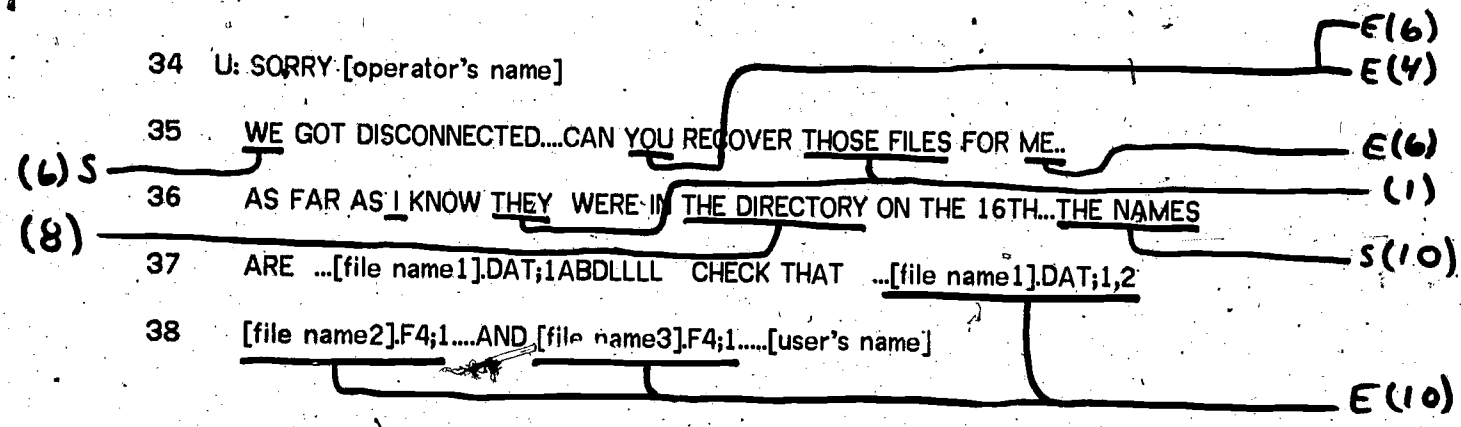


Figure 6-1b - Dialogue Annotated for Repeated Reference - Part b

7. *REQUESTS*

In the course of a dialogue, frequently one of the participants will communicate to his partner an expectation about the partner's subsequent behavior. We intend to include all such communications under the heading of ***Requests*** The Observer is asked to detect any occurrence of a speaker's indicating such an expectation, and to classify each of these utterances into one of five categories:

1. Questions -- request immediate, verbal response.
2. Orders -- request immediate, nonverbal response.
3. Directives -- request certain behavior in the future.
4. Rhetoricals -- look like ***Requests*** but are not.
5. Prohibitives -- request TO NOT DO something

Having done this, we ask the Observer to annotate the immediately following verbal response by the requester's partner. These annotations attempt to characterize the cooperative or uncooperative nature of this response.

OBSERVER'S INSTRUCTIONS FOR REQUESTS

We are interested in the general category of utterances which we call ***Requests***, by which a speaker communicates a specific expectation, request, or demand to the hearer. We are also interested in the cooperative or uncooperative nature of the hearer's next reply.

There are many different ways a speaker may express a ***Request***:

direct:	"Please pass the salt."
indirect:	"I could sure use some salt."

The speaker may ***Request*** behavior which is:

verbal:	"Is MTA0 write-enabled?"
nonverbal:	"Please mount JM11a on DTA3."
both:	"Do you have the time to retrieve my file X from archives?"

The behavior Requested may be:

immediate: "Attach the pump to the platform."
delayed: "Please call me when you arrive."

There is a set of utterances which look like Requests, but which both speaker and hearer know are not. The following are examples of this category of Requests:

"Why don't you go jump in the lake?"
"Who do you think you are?"
"Is the sky blue?"
"Go fight City Hall!"

Finally, there are Requests to not do or say something:

"Don't think about elephants."
"Whatever you do, please, don't throw me into the briar patch!"
"Don't put beans in your ears."
"Never tow your car while it is in gear."

We divide the set of possible Requests into the following groups:

1. QUESTIONS. For our purposes, "Question" refers to a much larger class of utterances than does the conventional, casual use of the word. By "Question" we mean any utterance by which the speaker communicates to the hearer an expectation, request, or demand for ***specific, immediate, verbal behavior.***

Verbal: The anticipated response is an utterance.

Immediate: The hearer is to initiate his compliance commencing with his very next speech act.

Specific: The expression is sufficiently detailed that it is potentially fulfillable by a single, appropriate response. ("How do I take a square root?" is specific; "Help me, I'm lost!" is not.)

2. ORDERS. By "Order" we mean any utterance by which the speaker communicates to the hearer an expectation, request, or demand for ***specific, immediate, nonverbal behavior.***

Immediate: The requested nonverbal behavior is to commence as soon as the hearer completes the process of comprehending the order.

Specific: The speaker communicates the belief that with this Order (and what preceded it) the hearer has sufficient details to enable him to perform the desired behavior.

3. DIRECTIVES. By "Directive" we mean an utterance by which the speaker communicates to the hearer an expectation, request, or demand for any sort of ***Delayed behavior.***

Delayed: the speaker communicates no sense of immediacy of response in his Request (including, of course, the case where a future time or condition is indicated explicitly).

4. RHETORICALS. A "rhetorical" is an utterance which has the form of a Request, but which, in the context of its use, is understood by all involved to represent something other than a mandate to the hearer.

5. PROHIBITIVES. Any Request to ***not engage in a particular behavior*** is a "Prohibitive." Note that this need not require a negative word ("Trespassers will be eaten!"). This class also includes the "don't do anything--don't go away--don't be impatient", implied by "Wait!" and its related forms.

Frequently a Request will be taken, by custom, to mean something different from (but possibly related to) the literal sense attributable to it. A simple example of this is the request "Do you have a match?" Clearly, those cases in which this is asked as a simple yes/no question would be regarded by the average speaker in our culture as atypical. For the purposes of the above classification, you will be asked to label a Request according to the principal force of the utterance -- the clearly intended and recognized meaning of the Request, independent of its surface form. In the case of the above example, "Do you have a match?" would be annotated as though it were "Please give me a match."

Any Request which falls into more than one of these categories should be annotated separately for each appropriate category.

Should you encounter an utterance which seems to satisfy the general definition of Requests described above, but which doesn't fit into any of these subcategories, then underline it and describe the new subcategory of which you feel it is a member.

Note that in nonpathological dialogue, any utterance may create a general expectation of relevance, continuity, etc. in the subsequent reply. The creation of this type of expectation is not intended to be included in our definition of Request. Thus, although in the sequence: A: "I don't feel well.", B: "Santa Monica.", B would not seem to have satisfied A's expectations, we would not regard A as having uttered a Request.

You are to annotate each of these categories of Request separately, according to the following directions.

Part 1 -- Questions.

Whenever you see a Question, enclose it in angle brackets, mark it with a "Q" and a unique number.

e.g.:

(Q5) ⟨Tell me your name.⟩

Do not rely on the presence or absence of question marks in the transcript? Is that clear. Remember that any contiguous region of an utterance that exhibits an expectation of an immediate, specific, verbal response falls into the category of a Question.

For each Question so annotated, examine the immediately following turn of the other person. We want you to separate two classes of turns: a) the turn provides (or begins to provide) the anticipated behavior, and b) everything else. If this turn contains one or more utterances which constitute some or all of the specifically requested verbal behavior, then annotate the turn with the Question label and a "+" (e.g.: "Q5+"). Otherwise, annotate the response with the same label and a "-" (e.g.: "Q5-").

Note that the turn may begin with verbal behavior outside that requested, ("Hmmm, gee, let me think, I don't know, well, yes!, I guess so."). As long as the recipient of the Question begins to provide the requested behavior prior to his partner's next turn, we want you to annotate this as "+". Do not attempt to annotate the intervening behavior with respect to this particular Question.

Be sure you distinguish between those responses which actually provide the desired behavior and those which (only) set in motion a chain of events which is expected to culminate in the desired behavior. An example of this would be a response which sought to clarify the question, thus indicating a willingness to answer it (but which was not itself initiating the answer).

For those responses which are marked "-", select one of the following descriptions which most closely captures the function performed by that response; append the description's label to the annotation (e.g.: "Q5-A3"). If none seems to apply, then invent and describe your own. If the response serves more than one of these functions, choose the one indicated first; if this distinction is not clear, indicate all the apparently simultaneous responses (e.g., "Q5-A1,A2").

A1. Requests clarification of meaning of Question or definition of a term in the Question.
A: How old is your grandmother?
B: Which one?

A2. Requests clarification of basis or cause for asking the Question.
A: Where were you last night?
B: Why do you ask?

A3. Challenges a premise or assumption in the Question.
A: Why did you go to the party without me?
B: What makes you think I was there?

(Note that A1 through A3 usually create new Questions.)

A4. Indicates that he may have or know the information but refuses to supply it.
A: Where did you get the $10,000?
B: I refuse to answer on the grounds ... .

A5. Dismisses Question as meaningless or otherwise unsuitable for answering.
A: Do you think the union will ask for more money?
B: I won't dignify that with a reply.

A6. Exhibits a promise, willingness or intention to reply to the Question, somehow, after a delay.
A: So what's the final figure?
B: I'll have that for you this afternoon.

A7. Disclaims knowledge of the requested information.
A: Can you tell me anything about LINK?
B: Sorry, I'm new here and don't know.

A8. Declines to supply the information.
A: How much do you make?
B: That's none of your business!

A9. Does not take up the Question in any way.
A: Young man, what are your intentions concerning my daughter?
B: Hasn't the weather been lovely lately?

Identify the region of the transcript which you used as the basis for your choice of labelling. Use double angle brackets, i.e., <<, >>. Label each bracket with the Question identifier, e.g. (Q5).

After you have read the whole transcript and identified all of the (Q)'s, go back and decide, for each one, whether the expected response was *ever* supplied in the transcript. If so, identify the region of the transcript in which it was supplied. Use the segment markers described in the general instructions, i.e., (=, (, [,), =). Label each marker with the Question identifier, e.g., (Q5). If there are substantial gaps (a sentence or more) in the region in which the response is supplied, use several sets of marked brackets rather than one set. For each Question, label the rightmost] or =), with the word "Partial" if only part of the requested verbal behavior was exhibited.

It may appear that the participants end up with different views on whether the requested behavior was supplied, or whether it was supplied in full. If so, then label your brackets with the participant's identifier when indicating where the information was supplied, e.g.,](Q5 for George). Use two sets of brackets if necessary.

Part 2 -- Orders.

A contiguous region of an utterance that exhibits an expectation of an immediate, specific, nonverbal response, falls into the category of an Order. Whenever you see an Order, enclose it in angle brackets and mark it with "O" and a unique number, as above.

For each Order so annotated, examine the immediately following turn. We want you to separate two classes of responses: a) compliant and b) everything else. There are three kinds of compliant responses:

C1. Asserts completion of requested activity.
A: Now attach the second bolt.
B: I've already attached all of them.

C2. Asserts current performance of requested activity.
A: Next, empty the number two tank.
B: I'm already doing that.

C3. Asserts willingness, ability, and intention to initiate requested activity forthwith.
A: Private, do you think you could find me a cup of coffee?
B: Yes sir! Right away sir!

Label the compliant response with the Order label and a "+" followed by the type of response (e.g.: "(O5+C2)"). If you feel that there has occurred a compliant response which does not fall into one of the above categories, feel free to invent a new category of your own. For those responses which are other than compliant, annotate them with the Order label and a "-", as above.

For those responses which are marked "-", select one of the following descriptions which most closely captures the function performed by that response; append the description's label to the annotation (e.g.:"(05-R3)"). If none seems to apply, then invent and describe your own.

R1. Requests clarification of meaning of Order.
A: Position the drain towards the front.
B: Which side is the front?

R2. Requests clarification of basis or cause for giving the Order.
A: Glue the bottom before the top.
B: What will that buy us?

R3. Challenges a premise or assumption in the Order.
A: Set up a meeting with Jones this afternoon.
B: Have you forgotten? He's on vacation.

(Note that A1 through A3 usually create new Questions.)

R4. Indicates that he may be able to perform the Order but refuses to perform it.
A: Loan me $100 until payday, if you have it.
B: I could find it, but I don't lend money to anyone.

R5. Dismisses Order as meaningless or otherwise unsuitable for performance.
A: Will you give me a ride to the station?
B: Sorry I can't, my car's in the shop.

R6. Exhibits a promise, willingness or intention to perform the Order after a delay.
A: Johnny, go cut the lawn.
B: I'll do it this afternoon.

R7. Disclaims ability to perform the Order.
A: Please mount my tape on drive 1.
B: Sorry, I'm new -- don't know how.

R8. Declines to perform the Order.
A: Subscribe now and save 20%!
B: No thanks, I'm really not interested.

R9. Does not take up the Order in any way.
A: Point the camera towards the meter.
B: How many washers are there supposed to be?

As above, identify the region of the transcript which you used as the basis for your choice of labelling. Use double angle brackets, i.e., <<, >>. Label each bracket with the Order identifier, e.g., (O5).

After you have read the whole transcript and identified all of the (O)'s, go back and decide, for each one, whether it was responded to anywhere else in the transcript. If so, identify the region of the transcript which constituted the response. Use the segment markers described in the general instructions, i.e., (=, (, [,],), =). Label each marker with the Order identifier, e.g., (O5).

For each Order, label the rightmost] or =), with the word "Partial" if only part of the requested behavior seems to have been provided. If he tried but did not succeed, that's "Partial," too.

Part 3 -- Directives.

Any contiguous region of an utterance which exhibits an expectation for specific behavior in the future falls into the category of a Directive. Whenever you see a Directive, enclose it with angle brackets, mark it with a "D" and a unique number, as above.

A Directive has the sense "Perform behavior X at some time, or under some circumstance, in the future." Treat each Directive as though it said "I would like for you to perform behavior X ... . Do you assent/commit, now, to behaving as indicated, at the appointed time or under the requested circumstances?".

Thus, a Directive combines the functions of a Question ("Tell me that you will ... ") with those of an Order (" ... perform a certain behavior."). With this in mind, the response for each Directive is to be annotated according to the directions for Questions (i.e., responding to a request for an immediate, verbal act). The exception to this is whenever the response is like the C1, C2, or C3 responses to Orders. Treat the requested action in the present tense. In this case, annotate as a "+" response to a Directive.

As above, identify the region of the transcript which you used as the basis for your choice of labelling. Use double angle brackets, i.e., <<, >>. Label each bracket with the Directive identifier, e.g., (D5).

After you have read the whole transcript and identified all of the (D)'s, go back and decide, for each one, whether it was responded to anywhere else in

the transcript. If so, identify the region of the transcript which constituted the response. Use the segment markers described in the general instructions, i.e., (=, (,.[,],), =). Label each marker with the Directive identifier, e.g., (D5).

For each Directive, label the rightmost] or =), with the word "Partial" if only part of the requested behavior seems to have been provided. If he tried but did not succeed, that's "Partial", too.

Parts 4 and 5 -- Rhetoricals and Prohibitives.

Whenever you see a Rhetorical or a Prohibitive, enclose it in angle brackets and mark it with a "R" or "P", respectively, and a unique number, as above. Do not attempt to annotate the immediate response to these Requests.

After you have read the whole transcript and identified all of the R's and P's, go back and decide, for each one, whether the corresponding (unexpected) behavior was ever indicated in the transcript. If so, identify the region of the transcript in which it indicated. Use the segment markers described in the general instructions, i.e, (=, (,.[,],), =). Label each marker with the Rhetorical or Prohibitive identifier, (e.g., R5 or P5). If there are substantial gaps (a sentence or more) in the region in which the response is supplied, use several sets of marked brackets rather than one set.

Final, general instructions.

If you observe any place where one participant seems clearly to misunderstand his partner's ***Request***, underline the passages which led you to conclude this and summarize, in your own words, the nature of the misunderstanding.

If you find a ***Request*** that seems to you to be a repetition, perhaps with additional detail, of a previous question posed by the same speaker, indicate this by noting, next to the label for the new ***Request***, "=" followed by the old ***Request***'s label (e.g., "=Q5").

You should be particularly careful with utterances which appear to be ***Requests*** but which, in context, are only presented as descriptions of behavior rather than a ***Request*** to perform the behavior. For example:

A: How can I see better in here?
B: Turn on your light, stupid!
A: Oh, thanks.

The point of this example is that B is conveying to A the information he requested, not necessarily asking A to do anything at all. For these cases, the apparent "***Request***" (by B) is simply to be ignored.

Figure 7-1 indicates a segment of dialogue annotated according to these directions.

(Q1) 26 U: OK [operator's name]<CAN YOU PEOPLE GO INTO ANY DIRECTORY AND DELETE FILES WHEN YOU

27 NOTICE THAT THEY ARE NOT BEING USED.> [user's name]

(Q1-A7) 28 <<I AM REALLY NOT SURE, HOWEVER I KNOW THAT THAT WOULD NOT BE

29 RESPONSIBILITY WE WOULD TAKE UPON OURSELVES TO JUDGE WHETHER OR

30 NOT YOU WANT YOUR FILES OR NOT.>>

31 [outside interruption here]

(Q2) 32 O: [user's name]<ARE YOU THERE?>

33 LINK FROM [user's name], JOB 21, TTY 22

34 U: SORRY [operator's name]

(O1) 35 WE GOT DISCONNECTED.<CAN YOU RECOVER THOSE FILES FOR ME..>

36 AS FAR AS I KNOW THEY WERE IN THE DIRECTORY ON THE 16TH...THE NAMES

37 ARE ...[file name1].DAT;1ABDLLLL CHECK THAT ...[file name1].DAT;1,2

38 [file name2].F4;1....AND [file name3].F4;1.....[user's name]

(P1)

(O1-R6) 39 O:<<OK HOLD ON JUST A MINUTE>AND I WILL TRY TO FIND THEM>>

40 U: RIGHT

41 [outside interruption here]

Figure 7-1a -. Dialogue Annotated for Requests - Part a

(Q3) 42 O: ⟨ARE YOU STILL THERE?⟩

(Q3+) 43 U: ⟪RIGHT⟫ [operator's name]

(Q4) 44 O: OK YES I HAVE FOUND THE FILES YOU ARE CONCERNED ABOUT ⟨ON THAT

45 FIRST ONE I ASSUME THE ABDLLLL WAS AN ERROR RIGHT?⟩

(Q4+) 46 U: ⟪RIGHT⟫ THAT SHOULD HAVE BEEN [file name1].DAT;1 AND 2.

47 O: OK THEY ARE HERE

48 [outside interruption here]

(Q5) 49 O: ⟨ARE YOU THERE?⟩

(Q5+) 50 U: ⟪RIGHT⟫

51 O: OK I HAVE FOUND THE FILES YOU WANT

52 [outside interruption here]

(O1) 53 O: OK I HAVE FOUND THE FILES YOU WANT [I WILL RETRIEVE THOSE FOR

54 YOU ALSO] OK

55 U: GREAT.

Figure 7-1b - Dialogue Annotated for Requests - Part b

8. EXPRESSION OF COMPREHENSION

In most dialogues, each participant wants the other to comprehend what he says. Since each person knows this, the participants often say things that serve to inform the other that some part of what was just said was comprehended, or not understood, or partially understood. Sometimes there are parts of the dialogue that perform only this function, and sometimes these parts do this while conveying other content. These are all aspects of Expressions of Comprehension that we want Observers to comment on.

There are two components of this mechanism for expressing Comprehension: there is the Expression itself, and there is the part of the dialogue being referred to. In many cases, this second part is the other person's previous turn. We want the Observer to comment on both of these parts.

We have found that these Expressions occur frequently in two kinds of situations: those with noisy channels and those where there is a high cost for misunderstanding. We expect to find them whenever their cost to the speaker is lower than the cost of the expected reduction in misunderstanding.

Here are the instructions for the Observer.

OBSERVER'S INSTRUCTIONS FOR EXPRESSION OF COMPREHENSION

We are interested in the Expression by a speaker that he comprehends previous dialogue. There are various ways in which this Expression may occur:

Three Explicit Examples:

"OK, That's clear."

"Sure."

A: "Is that clear?"
B: "Yeah."

An Implicit Example:

A: "...so that proves the lemma."
B: "Using that same method, I can prove the second one."

Our idea of comprehension here covers both hearing what is said and understanding it. We are interested in the hearer's Expression of his view of his own comprehension -- the confidence dimension rather than the correctness dimension. (Thus these instructions will not serve to identify cases in which the hearer misconstrues but does not know that he misconstrues.)

People in dialogue indicate various degrees or scopes of comprehension of what the other party has said. Using the directions below, we want you to identify places where someone is indicating his state of comprehension of what has gone before, and to note the degree and scope of that comprehension.

Three kinds of indication of comprehension will be noted: Positive Comprehension, Noncomprehension, and Selective Comprehension.

Positive Comprehension

Scan the transcript for explicit and implicit expressions of positive comprehension. Mark each one with angle brackets , < and >, and in parentheses the letters PC (for "positive comprehension") and a unique number, e.g., (PC5).

To qualify as an expression of positive comprehension, a region must either explicitly express comprehension or implicitly indicate this by specific dependence on what was said before. The lack of an Expression of noncomprehension is not sufficient evidence for implicit positive

comprehension. Instead, an utterance must rely on some specific knowledge expressed previously by the other person.

Some of these Expressions will not refer specifically to any definite part of the dialogue.

Others will clearly refer to some particular previous part. For the latter, mark the part(s) with the segment markers described in the general instructions, e.g. (=, (, [,],) and =). Mark each of these markers with the PC number, e.g. (PC5) which it is indicating the scope of. If the scope of the indication is exactly the previous comment by the other speaker, then a slash with the identifying number may be used instead of the scope brackets to indicate this; e.g. (PC5)/ .

Multiple indications: Often people will indicate comprehension in more than one way (according to our notation). For example, they will commonly give an explicit indication such as "Got that." followed by an implicit indication such as use of the comprehended material.

RULE: Annotate the earliest indication. The scope of the indication should be the largest contiguous scope which is either all Primary or all Nonprimary according to the directions given below.

Use this rule for all the comprehension categories.

The comprehension may be expressed to various degrees:

P1. Indefinite degree: Expressions like "Keep going" and "I guess so" indicate that some comprehension has occurred.

P2. Satisfactory or Substantially Complete Comprehension: Expressions like "OK" and "Sure" sometimes indicate this more complete level. Direct repetition without significant change of meaning is also used to indicate this level of comprehension. (But sometimes it indicates selective comprehension instead.) Expression may include approval or agreement or consent.

Mark each Expression of Comprehension with an indication of the degree of Comprehension expressed. Use one of the P numbers above, or assign a new P number of your own, and supply a descriptive phrase which tells qualitatively the degree of comprehension that was expressed. If you use your own phrase, try to express its rank relative to the phrases above, e.g., more than P1 but less than P2

Noncomprehension

Next, scan the transcript for explicit or implicit indications of noncomprehension. Phrases like "Huh?" and "Say that again" may indicate noncomprehension. Mark each such indication with NC (for "noncomprehension") and a unique number, as above, and indicate the scope of the noncomprehended region, if known, and its partialness, if it is clearly partially noncomprehended. Mark them in a manner directly analogous to the above directions for positive comprehension.

Expressing a doubt about one's own comprehension is a variety of expressing noncomprehension. Even when in fact all was well comprehended, an Expression of doubt should be marked as indicating noncomprehension.

Noncomprehension may be expressed to various degrees:

N1. Indefinite degree of noncomprehension.

N2. Substantially complete noncomprehension.

Selective Comprehension

People sometimes indicate what they have or have not comprehended. Where they do so, mark the region with angle brackets and a unique SPC number or SNC number (for "Selective Positive Comprehension" and "Selective Noncomprehension" respectively), as is done for PC numbers above. Use a new SPC or SNC number for each selected item. To qualify as being selective comprehension, there must be some indication that there are differences in how well the various candidates for comprehension have been received.

Indicate the region for which the comprehension or noncomprehension is being indicated, as above.

If you identify selective positive comprehension within a region which you also mark as having noncomprehension, this will be interpreted as noncomprehension of all except the selected item(s). Similarly, you can show comprehension of all except selected items with a PC and an SNC. for each particular thing identified selectively as comprehended.

Use the phrase marks (such as P1 and N2) from the Positive Comprehension and Noncomprehension directions to indicate categories of SPC and SNC.

The Primary/Nonprimary Distinction

For each indication of comprehension that you identified, decide whether it is used by the hearer primarily to indicate his comprehension, or (alternatively) whether indication of comprehension was performed along with some other primary function. Mark the segment identifying number with ++ if indication of comprehension is the primary function of the segment, and with -- if it is not. So for example "huh?" might be marked as (NC5)++, showing that it was identified as an indication of noncomprehension, and that the primary function of the segment was to indicate the speaker's lack of comprehension.

Use of the material comprehended, such as answering a question, is a common nonprimary indication of comprehension. Simple Expression of comprehension with approval or agreement or consent should be treated as a primary indication of comprehension.

People sometimes indicate their state of comprehension by telling the other person what to do about that state. Expressions like "Keep going" and "Say that again?" are used in this way. A speaker's indication of his state of comprehension in any of these ways should be marked with ++.

All of these indicators of comprehension deal with a person expressing something about his own state. Do not use this notation for cases in which one person is indicating that the other one is confused, has failed to comprehend, has comprehended well, or something similar.

Notation summary:

<, >	identify place where comprehension was expressed
(=, (, [,],), =)	identify place which Expression refers to
PC	segment expressing Positive Comprehension
NC	segment expressing Noncomprehension
SPC	segment expressing Selective Positive Comprehension
SNC	segment expressing Selective Noncomprehension
++	expression of comprehension is primary function
--	expression of comprehension is not primary function
/	(abbreviation) expresses comprehension of the previous comment by the other speaker

Figure 8-1 shows an example of applying these instructions to a sample transcript.

(PC1)

26 U: OK [operator's name]. [CAN YOU PEOPLE GO INTO ANY DIRECTORY AND DELETE FILES WHEN YOU

27 NOTICE THAT THEY ARE NOT BEING USED.] (PC1) [user's name]

(PC1) P2--

28 O: I AM REALLY NOT SURE) HOWEVER I KNOW THAT THAT WOULD NOT BE

29 RESPONSIBILITY WE WOULD TAKE UPON OURSELVES TO JUDGE WHETHER OR

30 NOT YOU WANT YOUR FILES OR NOT.

31 [outside interruption here]

32 O: [user's name], ARE YOU THERE?

33 LINK FROM [user's name], JOB 21, TTY 22

(NC2) / N2++

34 U: SORRY [operator's name]

35 WE GOT DISCONNECTED. CAN YOU RECOVER THOSE FILES FOR ME.

36 AS FAR AS I KNOW THEY WERE IN THE DIRECTORY ON THE 16TH...THE NAMES

37 ARE ...[file name1].DAT;1ABDLLLL CHECK THAT ...[file name1].DAT;1 2

38 [file name2].F4;1....AND [file name3].F4;1.....[user's name]

(PC3) / P2++

39 O: OK) HOLD ON JUST A MINUTE AND I WILL TRY TO FIND THEM

Figure 8-1 - Dialogue Annotated for Expression of Comprehension

9. SIMILAR EXPRESSIONS -- TWO KINDS

OVERVIEW

In modeling communication we particularly need means for judging how well the models we build perform comprehension or understanding. Clues about what the understanding of dialogue depends on are also vital. Developing data on which those judgments can be made is an essential step toward making evaluations of comprehension operational.

The Similar Expressions category is intended to provide a basis for knowing whether a model is extracting too much or too little or the wrong things from particular items in a dialogue transcript. It is also intended to provide very selective information about the dependencies of interpretation that arise between parts of an ongoing dialogue.

We are exploring this with two similar, but independent, experiments. The first of these calls for the Observer to generate alternative expressions which he feels would accomplish the same function as existing expressions from a real dialogue. In this case, he is given the full context of the surrounding dialogue. The second approach is significantly more complicated, involving the generation of these alternate expressions, but this time without the surrounding dialogue. The Observer is then given the surrounding dialogue and asked to separate the proposed alternatives into those which will work in the dialogue and those which will not.

The details of these two experiments will be spelled out separately below. However, the first step in each is the common one of segmenting the given dialogue into "units." This will be detailed first, with the understanding that it is to be taken as the initial step in each of the processes described subsequently.

Instructions to Unit Divider

You will be given a dialogue. Divide each turn into "units," where a unit is either a simple sentence, something that functions like a simple sentence, or the whole turn. Do not feel bound to observe the punctuation you find in the transcript it is likely to be inaccurate. The unit boundaries should reflect a single, coherent communication having approximately the "completeness" of a simple English sentence. In no case, however, is a unit to be larger than a whole turn. Assign a unique number to each unit.

Similar Expressions In Context

We would like to compare different ways of saying something, i.e., ways that are reasonably construed as being equivalent in communication effect. We will describe such expressions as Similar and intend the term to have a specific technical meaning.* By comparing the effects on the model of a dialogue item with the corresponding effects of the Similar Expressions, we can see whether the models are creating unnecessary differences or missing important common effects.

OBSERVER'S INSTRUCTIONS FOR SIMILAR EXPRESSIONS IN CONTEXT

We are interested in identifying different expressions that would seem to have the same effect in a specific dialogue.

You will receive a transcript of the dialogue that has been divided into "units" (according to the directions, above, for Unit Dividers). For each unit, compose one or more Similar Expressions of each unit, consistent with the following constraint:

If this new expression had occurred instead of the original unit, it would have been acceptable to the speaker for the purposes he had in mind. You should base your judgment only on the dialogue which precedes the utterance in question, not on any dialogue which follows it.

We encourage you to use different words and styles of expression than appear in the given units. These Similar Expressions constitute the results from the Observer, for this experiment.

EXAMPLE OF SIMILAR EXPRESSION IN CONTEXT

STANDARD:
Can you recover those files for me?

SIMILAR EXPRESSIONS IN CONTEXT:
Is there any way to get my files back?
Please restore these files if possible.

---

*Two expressions are Similar if, in your opinion, the speaker would be willing to use the second at the point where he actually used the first.

We attempt to express this "definition" a little more clearly with the following hypothetical situation: Imagine a dialogue between a speaker and a hearer who did not share a common language, but who had translators available. Imagine further that in order to insure accuracy of communication, the speaker speaks to Translator1 (only) who translates what he hears and communicates this translation to a second translator who cannot hear the original utterance. The second translator then retranslates the utterance from Translator1 back into the first language for the benefit of the speaker. Thus, the speaker is afforded the opportunity to hear his own words repeated back to him, after undergoing two passes of translation. At this point he has the option of accepting the echoed utterance as adequate, or rephrasing it and trying again.

Within the framework of this hypothetical situation, two expressions are to be considered Similar if, after saying the first, and subsequently hearing the second echoed through the second translator, the original speaker would be willing to permit the differences to go unchallenged -- that is, he accepts the different version as sufficiently adequate for his original purposes that it does not need revision.

Similar Expressions Out Of Context

There are many different sources of information that the interpretation of dialogue text typically depends on. They include

- Situational Information- including the circumstances of the dialogue, the identities and cultural roles of the participants;
- Dialect and Jargon Information- The local conventions of language usage that were in effect at the beginning of the transcript.
- Context Information- The part of the transcript that precedes the element being interpreted.

Interpretation depends on each of these in the sense that for a given item (sentence, word, or other) variation in each of these related information sources can produce variation in the interpretation of that item in dialogue. (Notice that the term "Context" is being used in a specific, narrow and examinable way. It is not a catchall term for "other effects.")

Because we are dealing explicitly with the effects of each item in the transcript, identifying effects of context is particularly important. These directions are designed to identify these effects and discriminate them from the others.

The basic scheme for identifying context effects is sketched in Figure 9-1. Instructions for each step appear below. The point of the process is to generate expressions that are judged Similar in the out-of-context equivalence judgment, and then to judge the acceptability of these with the in-context equivalence judgment. THose that are unacceptable in this second stage must be so because of some effect of

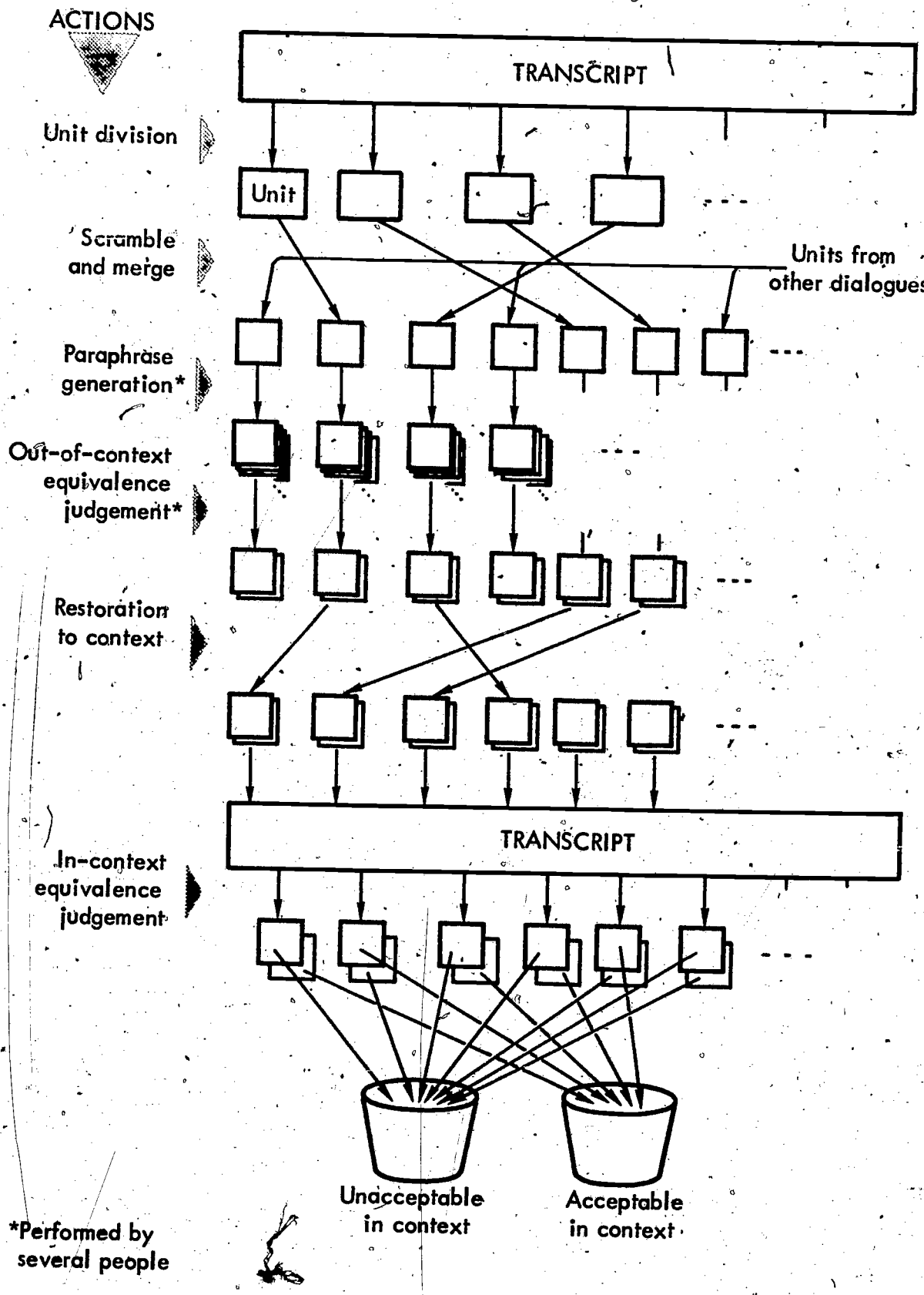


Figure 9-1 - Information Development for Similar Expressions Out of Context

context. We ask the Observer in the second judgment to describe factors that make such items unacceptable. We can then evaluate a model by examining whether a corresponding factor (difference of effect) arises when the proposed Similar Expression, rather than the original item, is read by the model at the appropriate moment.

OBSERVER'S INSTRUCTIONS FOR SIMILAR EXPRESSIONS OUT OF CONTEXT

We are interested in the effects of the preceding dialogue on the interpretation of a turn in that dialogue. In order to study these effects, we have devised a multistep procedure which makes use of your observations. It starts with a dialogue transcript, isolates pieces of some turns from their surrounding dialogue, identifies alternate, Similar Expressions, not knowing the context in which the unit originally occurred, and finally implants these proposed alternates in the original dialogue at the point where the original piece had been, for judgment on how well or poorly these interpretations fit.

The Similar Expressions are prepared out of context (i.e., without knowing what preceded or followed the original unit). The ones that turn out later to fit back into the context are used to develop methods of interpreting concepts that can be expressed in more than one way. The Similar Expressions that are later judged not to fit are used to identify kinds of differences between various ways of saying things, leading to development of methods for being responsive to those differences. Both the Similar Expressions that fit into the context and those that fail are valuable.

(Several Observers will perform the unit division task on different dialogues. For each Observer in the next stage, we will select some units from each of these dialogues and combine them into a set having no particular order or relation between the units as they are presented.)

SECTION 1: Instructions to Generator of Similar Expressions

We are interested in different ways of saying something which have the same potential effect. We want you to produce alternatives to some particular units that have occurred in actual dialogue.

You will receive a set of units. For each of these, we want you to write one or more new units (Similar Expressions) that are Similar to the original unit in the following way: In your judgment, the new expressions would be regarded by the speaker as having about the same effect as his original unit in some common, ordinary circumstance.

We encourage you to use different words and styles of expression than appear in the original units. Only the effect which you think the speaker would expect needs to remain the same.

Generating these Similar Expressions might be easier if you proceed in the following steps:

1. Read the unit.
2. Imagine a situation in which the unit could occur.
3. Imagine the effect intended by the speaker.
4. Invent a different unit which would have the same effect in the situation which you imagined. Write this new unit down.
5. Repeat steps 2 to 4 for as many different situations as come readily to mind. (Note: "different" is to be interpreted as meaning situations which would suggest a different set of admissible Similar Expressions.)

You should try to imagine ordinary situations rather than outlandish ones, since the expressions which you create will be judged by other Observers on the basis of whether they would have the same effect in some ordinary circumstances.

For example, in attempting to generate Similar Expressions for the unit: "yes," the following are two common contexts and one we regard as outlandish, each followed by a set of Similar Expressions appropriate to that context:

1. ("Is *he* your boss?") He sure is./That he is./He's the one.
2. ("Do you want to go to the movie?") You bet./Any time you're ready./fine.
3. ("What's a three letter word meaning affirmative?") Oui.

(For each unit, we will assemble a list of all Similar Expressions generated in all contexts by all Observers. The next step Observers will be given several sets of these expressions, each with the original unit that inspired that set. He still will not know the context in which the original unit occurred.)

SECTION 2: Instructions to judge of Similar Expressions (out of context).

We are interested in whether various ways of saying something are likely to have the same effect.

You will receive a collection of units, divided into groups. One in each group will be called the Standard Unit, and the others will be called the Comparison Units. You are to judge each Comparison Unit according to the following criterion: In your judgment, would the effect of the Comparison Unit be acceptable to the speaker of the Standard Unit in some ordinary circumstances?

Mark each Comparison Unit with one of the symbols: "+", "-", or "*" for Acceptable, Unacceptable, or Unclear.

(The Observer for the final step will be given one of the original transcripts with the units numbered. For each numbered unit, he will also be provided with a set of those Similar Expressions for that unit which received a "+" in the prior step's grading.)

SECTION 3: Instructions to judge of Similar Expressions (in context).

We are interested in how the effect of what is said is related to the context in which it is said. We want you to make some judgments about whether or not certain Similar Expressions for a unit can actually take the place of that unit in context, and, if not, why not.

Materials:

You will receive a transcript of an actual dialogue, as it occurred, divided into numbered units. For each numbered unit, you will also receive a set of possible alternative expressions for it.

Commentary:

For each one of the Similar Expressions given to you, make two kinds of annotation:

1. If this expression had occurred instead of the original, would it have been acceptable to the speaker of the original? Mark it with either "+", "-" or "*", for Yes, No, or Unclear. You should project the acceptability of the turn based only on the part of the dialogue that precedes the unit, not on anything which follows.

2. If you answer "-", then describe very briefly one important difference between the original and the alleged Similar Expression, from the point of view of the speaker. If the expression is incomprehensible in this context, mark it with "X", meaning Incomprehensible (even if it might be meaningful in other contexts).

These annotations constitute the results of the second experiment.

- EXAMPLE OF SIMILAR EXPRESSIONS GENERATED OUT OF CONTEXT

STANDARD UNIT:
As far as I know they were in the directory on the 16th.

RESULTS FROM SECTION 1:

1. They weren't gone before the 16th.
2. I think their phone was listed on the 16th.
3. My files were still there the 16th.
4. They seem to have been in the directory on the 16th try.

RESULTS FROM SECTION 2:

1. +
2. −
3. +
4. +

CONTEXT PRECEDING STANDARD UNIT:

- Can you recover those files for me?

RESULTS FROM SECTION 3:

1. +
3. +
4. -(Number clearly refers to date, not number of "tries".)

10. TOPIC STRUCTURE

The participants in a conversation often continue to talk about a particular set of concepts for several turns. This related set of ideas constitutes a *Topic* of discussion. Topics may proceed for a long time, as in a long telephone discussion of vacation plans, or may proceed for only two short turns, as in a simple question about the time its the answer. Topics may be general or specific, and the participants may be discussing more than one Topic at a time. Topic is a unit based on the content being discussed rather than the forms being used. Some parts of the discussion may be part of no clearly distinct Topic - for example, the negotiation of what to discuss next. The existence and nature of units of discourse larger than a single sentence is one of the important issues being currently pursued by a number of researchers. The Topic annotation described here may shed light on the currently hypothesized supersentential structures, such as "frames" (Minsky, 1974), "scripts" (Schank & Abelson, 1975), and "story grammars" (Rumelhart, 1975).

Here is the current set of instructions for annotating Topics in dialogue:

OBSERVER'S INSTRUCTIONS FOR CODING TOPICS

In most dialogues there are one or more Topics which are being discussed. A Topic is a subject discussed in one or more turns of the dialogue.

We want you to go through the dialogue, marking those segments in which a Topic is being discussed. Annotate the beginning and ends of these segments with the Segment Markers, described above. Note that segments may contain other embedded segments.

Do this for each person, separately in the dialogue. Also, draw a line from the initial marker for each Topic to the margin (left for speaker A; right for speaker B) and write a brief descriptive title for the Topic. Use this same title whenever the same Topic reappears, in either speaker's utterances. Notice that not every part of the transcript has to be part of a Topic.

When you have finished annotating the Topics of the dialogue, list any Topics still open at the end of the transcript. Also list any Topics that were already open when the part of the dialogue you have annotated started. To check these, make sure that every Topic marked as opened is either marked as being closed or listed as still open. Do the same for every annotation of Topic closing.

Figure 10-1 is an example of an annotation resulting from the application of these instructions to a piece of dialogue.

53 O: OK I HAVE FOUND THE FILES YOU WANT I WILL RETRIEVE THOSE FOR

54 YOU ALSO, OK

Retrieving Files

Retrieving Files

55 U: GREAT WAS THAT [name3] YOU WERE TALKING ABOUT?

People at the Computer Center

People at the computer center

56 O: NO, IT IS [name4], HOWEVER WHEN YOU SEND MESSAGE

57 SEND TO [name2]

58 U: RIGHT...IS [name5] STILL AR[outside interruption here]

59 IS [name5] STILL AROUND.

60 O: NO HE ISN'T. HE'S BEEN GONE FOR ABOUT THREE WEEKS NOW.

61 U: OK THANKS FOR YOUR HELP , [operator's name] HAVE A GOOD DAY [user's name] OUT.

People at the Computer Center

62 O: OK YOU DO THE SAME WILL DO THOSE FILES RIGHT AWAY. BYE

People at the computer center

63 U: THANKS AGAIN

64 BREAK

Topics already open at beginning	Topics not closed at end
O: Retrieving files	none
U: Retrieving files	

Figure 10-1 - Dialogue Annotated for Topic

SHARED TOPIC POST-PROCESSING

In our early directions, we attempted to have the Observer annotate *the* topic shared by the two participants. This led us into considerable difficulty, since the topics addressed by the individual participants were frequently "out of step" with each other in nontrivial ways. From our attempt to capture the notion of one common topic of a dialogue, we developed the concept of a particular participant's topic. Thus, the two participants may be talking about different topics at the same time, as often occurs at topic boundaries, when one starts a new topic before the other finishes talking about the old one. This notion of a individual's topic is imbedded in the instructions given above.

However, we found that we could derive from these observations of an individual's topic an annotation of shared topic (involving no additional subjective judgments), following the simple algorithms given below.

Post-Processing Procedure For Shared Topics

We will use the Observer's judgments of topic boundaries to determine the period during which the two participants share a topic of conversation.

The general intention embedded in the algorithm is as follows: A participant is sharing the topic currently being discussed by his partner unless the Observer says that he definitely is not.

The detailed blow-by-blow procedure is:

1) To mark the beginning of a shared topic:

   a. Find a shared topic (T) (a topic discussed by both participants and given the same name by the Observer).
   b. Find the place that each definitely starts talking about T (the inner starts).
   c. Call the earlier start SE; call the later start SL.
   d. For the person who spoke SL, (PL), find the earliest point in the speech of either party at which: (SE had already occurred) and (PL had not spoken anything which occurred prior to the outer start of topic T by PL).
   This is the Topic Sharing Begin (TSB).

2) To mark the end of a shared topic:

   a. For each shared topic (T) marked as beginning, above.
   b. Find the places where it definitely ends for each participant (the inner ends).
   c. Call the later end EL; call the earlier end EE.
   d. For the person who spoke EE, (PE), find the latest point at which (EL had not yet occurred) and (PE had not spoken anything which occurred after the outer end of topic T by PE).
   This is the Topic Sharing End (TSE).

3) If, after steps 1 and 2, it is discovered that the TSE precedes the TSB, then merely delete both marks.

4) The region of sharing of topic T runs from the TSB to the TSE.

The following examples are intended to illustrate, in an abstract form, the desired intent of these directions. The five examples represent two utterances each by A and B, alternating, in the course of which a shared topic is initiated. The segment markers within the sequences of A's and B's indicate their position within the dialogue dictated by the Topic Annotation directions. On the line immediately below this, the dashed line indicates the region of the dialogue for which the topic is deemed to be shared. (On

the belief that start and end instructions are symmetric, only starts are shown.)

1) [AAAAAAAAA [BBBBBBBBB AAAAAAAAAA BBBBBBBBBB

---------> ...

2) [AAAAAAAAA BBBB[BBBBB AAAAAAAAAA BBBBBBBBBB

---------> ...

3) AAAAAAAAAA BBBBBB(=BB AAAA[AAAAA BB(BBBBBB

---------> ...

4) AAAAA[AAAA BBB(=BB(BB AAAAAAAAAA BBBBBBBBBB

---------> ...

5) AAAAA[AAAA BBBBBBBBBB AAAAAAAAAA (=BB(BBBBB

---------> ...

In figure 10-2, we exhibit an example of the annotation that results from applying this procedure to the individual topic example given previously.

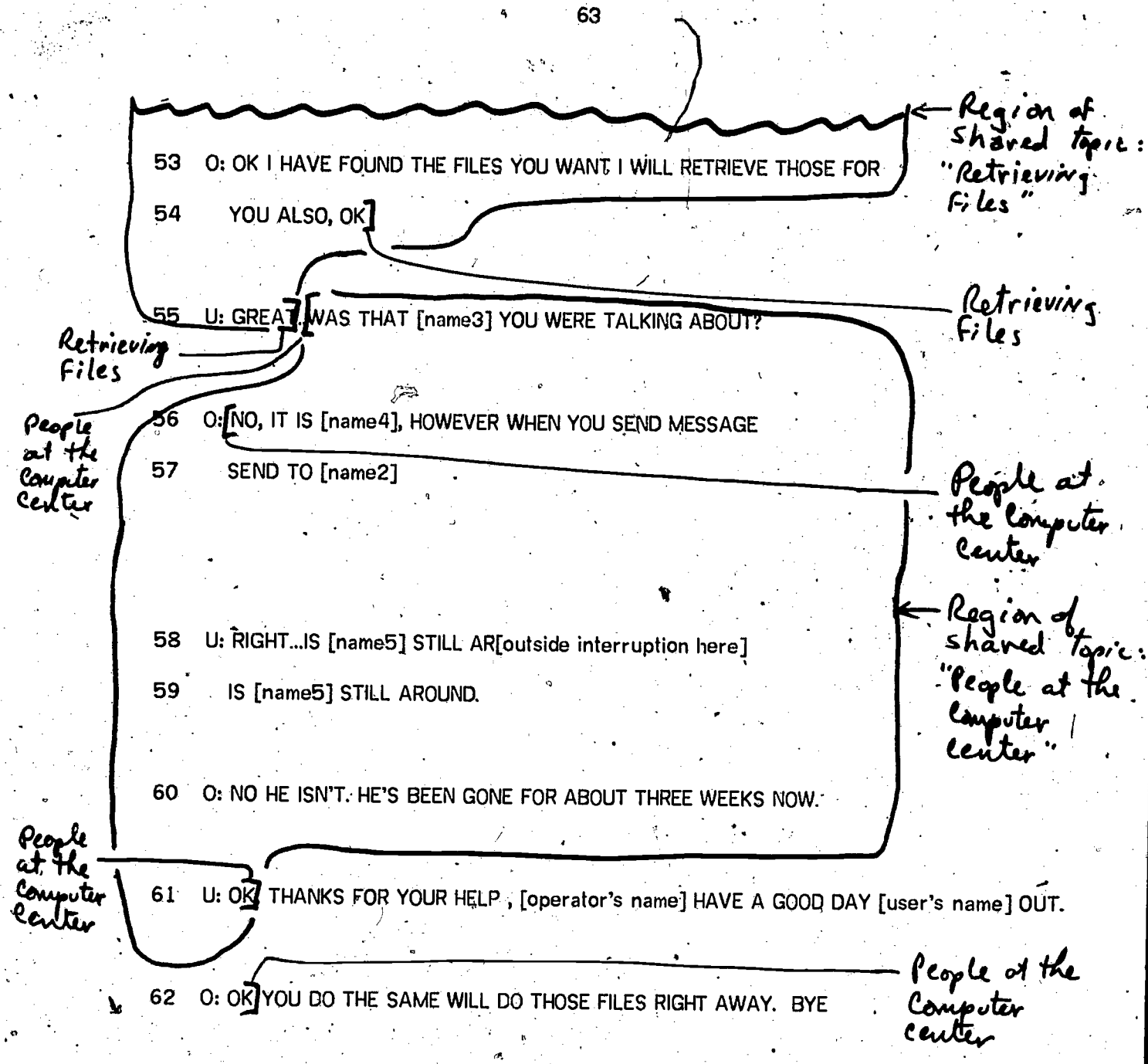


Figure 10-2 - Derived Shared-topic Annotations

11. INCREMENTAL AND PREREQUISITE INFORMATION

One of the most difficult aspects of a dialogue to grasp (at least for our modeling efforts) is what we loosely call the informational content. In order to somewhat confine our Observers, we have specifically limited their annotations on the information evidenced by the speakers, to that relevant to the topics, previously annotated. Since we felt any choice of formal notation for informational content would seriously bias the Observers, we have opted for simple English as the medium of expression for these observations.

Finally, we have asked the Observer to differentiate two kinds of information: that conveyed by the dialogue and that which was never expressed but must have been present for the dialogue to be understandable to the participants.

Although we are in full agreement concerning the importance of this particular category, we have been unable to achieve even a partial consensus on any proposed set of directions. We remain committed to the attempt to deal with this dimension but do not believe the current state of the directions merits being reported here.

12. CORRECTION ACTIONS

Every so often in the course of a dialogue, one of the participants becomes aware of a specific bit of misinformation which his partner seems to be operating with. When the speaker sets about to repair this unfortunate state of affairs, we say he is engaged in a Correction Action. Whenever we find one of these actions, we would like to identify:

1) Where did the correction occur?
2) What indicates that this is a correction?
3) What was being corrected?
4) How do we know what is being corrected?
5) How is the correct information indicated?
6) How could the correction have been avoided?
7) Did the recipient of the correction suggest it?
8) Was the corrector, himself, in error?

OBSERVER'S INSTRUCTIONS FOR CORRECTION ACTIONS

We are interested in how people make corrections during the progress of a dialogue. We want to identify corrections that occur, and eventually understand how they are accomplished and what their effects are.

For the purposes of this set of observations, we want to focus only on particular kinds of corrections, ignoring all others. The Correction Actions that we want you to identify and annotate must have the following two properties:

A. Retraction/Cancellation Property:

Something in the previous dialogue, which has been comprehended in some way by each participant, is retracted or cancelled by the further utterances of one of the participants.

Things retracted occur in a variety of forms. For example, the retraction may be effected by talking about retraction in some way, or it may occur as part of a substitution or revision.

The thing corrected and the correction may occur in the same turn; in this case it is sufficient to find that, in the absence of the correction, the recipient would probably have comprehended the part of the comment being corrected in another way.

The Correction Action may span several turns by the person doing the correcting, or it may be complete in a single turn.

B. Explicitness Property:

> There must be some explicit indication that a Correction was intended (e.g.: "No, I meant ... ", "Sorry, make that ... ", ...). So repetitions, clarifications, explanatory elaborations and restatements are not usually Correction Actions.

In our experience, Corrections having these two properties are relatively rare. Many dialogues do not contain any. On the other hand, it is possible to find dialogues in which they are relatively frequent.

The annotations are to be done in two stages, on separate copies of the transcript. The first stage deals with the regions which do the correcting; the second, with the regions being corrected.

ANNOTATIONS FOR CORRECTION ACTIONS -- first pass

A. CORRECTION REGION (CR)

Before you can annotate any of the details of a Correction, you must first ascertain that a Correction has occurred. When you find a region of the dialogue which satisfies the two requirements above (a Correction Region), bracket the region as you would a topic and assign the region a unique label: CR followed by a number (e.g. CR7). The Correction Region is to start with the first indication that a Correction is under way (this would include a "Correction Request" by the recipient, see below) and end when the Correction is no longer under discussion, for example,

A: ... out window three. [Correction on that! That's out window one.] (CR3)

Beyond this point, you should be aware that not all of the requested annotations will apply to each Correction. Likewise, as usual, it may not always be obvious which classifications apply or where. In either case, you are simply to ignore the inapplicable instructions or those whose application is not obvious.

Underline and label each of the regions described below with their two-letter code (e.g.: CP, CC, CQ, ...) followed by the number assigned to the corresponding Correction Region. You might find that some of these regions overlap or are even identical with others--that's perfectly all right.

B. CORRECTION FLAG (CF)

A consequence of the Explicitness Property is that there is almost always a part of the Correction Region (called the Correction Flag) in which the corrector signals that a Correction is taking place. (e.g. "Oops, what I meant was ... " , "Wrong-o!, the right answer is ....", ...) For each Correction Region you bracket, find its Correction Flag, for example,

A: ... out window three. [Correction on that! That's out window one.]
(CF8)

C. CORRECTION POINTER (CP)

Within the Correction Region, the corrector will use certain words to indicate what part of the preceding dialogue he wants to correct. This region (the Correction Pointer) will not represent any new or different information, for example,

A: ... out window three. [Correction on that! That's out window one.]
(CP3)

D. CORRECTION CONTENT (CC)

In addition to locating the Error Region with the Correction Pointer, the Correction Region will also indicate the nature of the substitution or revision. This novel information is the Correction Content, for example,

A: ... out window three. [Correction on that! That's out window one.]
(CC3)

E. CORRECTION REQUEST (CQ)

Occasionally, the recipient of an error will explicitly indicate his doubts that his partner indeed meant what he clearly said. (E.g.: "Did you mean ... or ...?", "Shouldn't that be ... ?", ...) If you see this happening, note this as a Correction Request. Note that this will then be the beginning of a Correction Region, for example,

A: ... out window three.
B: [Are you sure you mean three?
(CQ3)
A: Correction on that! That's out window one.]

Figure 12-1 gives an example of first-pass Correction Action annotation.

ANNOTATIONS FOR CORRECTION ACTIONS -- second pass

On the second copy of the dialogue, write the brackets and labels from the Correction Regions, as annotated in the first pass. Once this is done, proceed with annotating the regions described below.

A. ERROR REGION (ER)

In most cases, the corrector will be making his Correction as a modification of some prior region of the dialogue which contained the error. This part of the dialogue (the Error Region) may be in his own words, or his partner's. For each Correction Region, find the corresponding Error Region, for example,

A: ... out window three. [Correction on that! That's out window one.]
(ER3)

34 U: SORRY [operator's name]

35 WE GOT DISCONNECTED....CAN YOU RECOVER THOSE FILES FOR ME..

36 AS FAR AS I KNOW THEY WERE IN THE DIRECTORY ON THE 16TH...THE NAMES

37 ARE ...[file name1].DAT;1ABDLLLL [CHECK THAT ...[file name1].DAT;1,2]

38 [file name2].F4;1....AND [file name3].F4;1....[user's name]

(CR1) (CC1) (CF1) (CP1)

39 O: OK HOLD ON JUST A MINUTE AND I WILL TRY TO FIND THEM

40 U: RIGHT

41 [outside interruption here]

42 O: ARE YOU STILL THERE?

43 U: RIGHT [operator's name]

(CR2)

44 O: OK YES I HAVE FOUND THE FILES YOU ARE CONCERNED ABOUT [ON THAT

45 FIRST ONE I ASSUME THE ABDLLLL WAS AN ERROR RIGHT?

(CQ2)

46 U: RIGHT THAT SHOULD HAVE BEEN [file name1].DAT;1 AND 2.]

(CC2)

47 O: OK THEY ARE HERE

(CF2) (CP2)

Figure 12-1 - First-pass Correction Action Annotations

The error region should include not only the text which is actually corrected, but also that which is used by the Correction Pointer (see below) to locate the area to be corrected.

Sometimes, the corrector is not correcting preceding dialogue, but rather an erroneous concept which he believes the recipient holds. If you see this happening, describe this erroneous concept in the margin and treat it as the Error Region.

B. CORRECT ALTERNATIVE (CA)

Frequently it is easy to see how the error which was corrected could have been avoided in the first place. If possible, select a region of the dialogue (usually the Error Region) and rewrite it, using as nearly as you can the style of the speaker, so that, had the substitute been used, the net effect would have been the same as the actual utterance and subsequent Correction, for example,

A: ... out window three. [Correction on that! That's out window one.]

(CA3)
OUT WINDOW ONE

C. CORRECTOR ERROR (CE)

If you encounter a situation where, in your judgment, the corrector makes a Correction which, if successful, will have no net effect, indicate this with a STAR (*) in the margin, alongside the Correction Region, for example,

A: Turn off switches one, two and seven.
B: O.K., will do.
(*) A: [Check that, switch two should also be off.]
B: Roger, switch two off.

Figure 12-2 shows an example of second pass annotations for Correction Actions.

34 U: SORRY [operator's name]

35 WE GOT DISCONNECTED....CAN YOU RECOVER THOSE FILES FOR ME..

36 AS FAR AS I KNOW THEY WERE IN THE DIRECTORY ON THE 16TH...THE NAMES

37 ARE ...[file name1].DAT;1 ABDLLLL [CHECK THAT ...[file name1].DAT;1,2]

38 [file name2].F4;1....AND [file name3].F4;1.....[user's name]

(CR1)

(ER1), (ER2)

(CA1) = (CA2) = [file name1].DAT;1,2

39 O: OK HOLD ON JUST A MINUTE AND I WILL TRY TO FIND THEM

40 U: RIGHT

41 [outside interruption here]

42 O: ARE YOU STILL THERE?

43 U: RIGHT [operator's name]

44 O: OK YES I HAVE FOUND THE FILES YOU ARE CONCERNED ABOUT [ON THAT

45 FIRST ONE I ASSUME THE ABDLLLL WAS AN ERROR RIGHT?

(CR2)

46 U: RIGHT THAT SHOULD HAVE BEEN [file name1].DAT;1 AND 2.]

47 O: OK THEY ARE HERE

Figure 12-2 - Second-pass Correction Action Annotations

13. SUMMARY AND PLANS

The report above has described the very substantial progress that has been made on identifying suitable dialogue phenomena for this work. The observation methods are all in a state of incomplete development, which will necessarily persist for some time. Since they are not in a finished state, it is not timely to try to make them completely adequate for widespread use or formally assess their reliability.

The substantive content of the observational categories must be fitted to the modeling process in order to have a smoothly operating methodology. This fitting has not yet been done. Until it is, the basic definitions of the categories will not be stable. Therefore formal validation work or extensive documentation work on the observational methods is not appropriate at this time. The impact of later modeling activity would effectively cancel any immediate work of either kind.

On the other hand, there is value in exposing the methodology and observational categories to technical and personal points of view not represented among the authors.

We plan to construct process models for one or two very short dialogues immediately. The purpose of these models will be to sketch the form of the processes needed and determine major relationships between parts. This stage will necessarily involve a subset of the Observers' categories, relatively rough representation of some kinds of knowledge, and short dialogues.

As our modeling skill develops, we expect to deal with more phenomena per model, with longer dialogues, and with higher fidelity to the actual communication events.

Past experience on related problems indicates that such models often converge rapidly on some relatively effective processes, each episode of accounting for new data taking substantially less work than the previous one. When this convergence takes place, it is indicative of success. Also, ability to model rapidly makes it possible to build successively more ambitious models.

Different parts of these models will mature at different rates. It is to be expected that some processes will be ready for transfer into working man-machine systems relatively soon, and others much later. We plan to identify and report on those that appear to be ready for transfer on a continuing basis.

APPENDIX
FULL TEXTS OF THE DIALOGUE OF THE EXAMPLE

The examples used in the report are not character-by-character reproductions of the original typed dialogue. Some differences have been introduced for various reasons, including privacy, ease of use, and a desire to focus on certain phenomena. This section makes it possible to identify some of the differences by showing a text which more strongly resembles the source text, but with privacy preserved and genuine irrelevancies deleted. (The text which resembles the source is shown, first, then the cleaned-up version prepared for the Observer.)

We have chosen to use cleaned-up data in order to limit the diversity of phenomena which need to be dealt with at once. Because the differences introduced in this way seem distinct from the phenomena of central interest, is therefore a reasonable tactic to eliminate them in the immediate future, without in any way ruling them out of the scope of the general problem of understanding how communication works.

The transcript below represents communication between a computer operator and a user of the computer. Their terminals have been linked together by use of the LINK command, which causes each character which appears at either terminal to appear at the other as well. The transcript differs from the original only in the following ways:

1. Material deleted is noted by [square brackets]
2. "O" indicates the operator's comments, and "U" the user's.
3. Each turn by a speaker has been set off by a blank line and these labels.

The spelling, grammar, punctuation, etc. are those that originally occurred.

LINK FROM [user's name], JOB 21, TTY 3

U: HELLO ARE YOU THERE...GA

O: YES, GO AHEAD

U: A COUPLE A QUESTIONS....ARE THE FILES THAT ARE ON-DISK ARCHIVED EVERY WEEK/

O: YES

U: I XXX
IN ONE OF MY DIR. I HAD SOME D/ FILES IN ONE DAY AND THE NEXT THEY WERE GONE AND THEY WERE NOT ARCHIVED/

O: WELL, I AM REALLY NOT THAT FAMIL SURE OF HOW THE ARCHIEING OF THE FILES ARE DONE SO I WOULD SUGGEST THAT YOU SEND A MESS. TO [name2] IN REQ\QGARDS TO THIS PROBLEM. THE ARCHIE\EVEING IS DONE DURING THE SWING SHIFT SO IN ORDER THAT YOU DO DONT NOT GET MES MISLED BY SOMETHING I AM NOT CERTAIN ON I WOULD RATHER

YOU DISCUSS THIS WITH [name2]. HE IS OUT OF HIS OFFICE AT THIS TIME BUT YOU COULD SEND HIM A MESS AND I AM SURE HE S WILL GIVE YOU ALL THE INFO YOU WOULD LIKE. [operator's name]

U: OK [operator's name]. CAN YOUPEOPLE GO IN TO ANY DIR AND DEL V/ FILES WHEN YOU NOTICE THAT THEY ARE NOT BEING USED.. [user's name]

O: I AM REALLY NOT SURE, HOWEVER I KNOW THAT THAT WWOULD NOT BE RESPONSIBILITY WE WOULD TAKE UPON OURSELVES TO JUDGE WHETHER OR Y NOT YOU WANT YOUR FILES OR NOT.

[outside interruption here]

O: [user's name], ARE YOU THERE?

LINK FROM [user's name], JOB 21, TTY 22

U: SOER

RY [operator's name]

WE C/ GOT DIC/SCONNECTED....CAN YOU RECOVER THOSE FILES FOR ME..

AS FAR AS I KNOW THEY E/WR// WERE IN THE DIR ON THE 16TH...THE NAMES ARE ...[file name1].DAT;1ABDLLLL/////////// CHECK THAT ...[file name1].DAT;1,2 [file name2].F4;1....AND [file name3].F4;1.....[user's name]

39 O: OK HOLD ON JUST A MIN. AND I WILL TRY TO FIND THEM

40 U: RIGHT

41 [outside interruption here]

42 O: ARE YOU STILL THERE?

43 U: RIGHT [operator's name]

44 O: OK YES I HAVE FOUND THE FILES YOU ARE CONECERNED ABOUT, ON THAT
45 FIRST ONE A ASSUME THE ABDLLLL WAS AN ERROR RIGHT?

46 U: RIGHT THAT SHOULD HAVE BEEN [file name1].DAT;1 AND 2/

47 O: OK THEY ARE HERE

48 [outside interruption here]

49 O: ARE YOU THERE?

50 U: RIGHT

51 O: OK I HAVE FOUND THER\R Y\Y THE FILES YOU WANT

52 [outside interruption here]

53 O: OK II HAVE FOUND THE FILES YOU WANT I WILL RETRIEVE THOSE FOR

54 YOUXXX

55 YOU ALSO, OK

56 U: GRA/EAT..WAS THAT [name3] YOU WERE TALKING AV/BOUT?

57 O: NO, IT O\O IS [name4], HOWEVER WHEN YOU SEND MESS

58 SEND TO [name2]

59 U: RIGHT...IS [name5] STILL AR[outside interruption here]

60 IS [name5] STILL AROUND.

61 O: NO HE ISN'T. HE'S BEEN GONE FOR ABOUT THREE WEEKS NOW.

62 U: OK, THANX FOR YOUR HELP, [operator's name] HAVE A GOOD DAY [user's name] OUT.

63 O: OK YOU DO THE SAME WILL DO THOSE FILES RITHT AWAY. BYE

64 U: THANX AGAIN

65 BRA\AEAK

The transcript below is the cleaned up version.

10 LINK FROM [user's name], JOB 21, TTY 3

11 U: HELLO ARE YOU THERE

12 O: YES, GO AHEAD

13 U: A COUPLE A QUESTIONS....ARE THE FILES THAT ARE ON DISK ARCHIVED EVERY
14 WEEK?
15 O: YES

16 U: IN ONE OF MY DIRECTORIES I HAD SOME FILES IN ONE DAY AND THE NEXT THEY WERE
17 GONE AND THEY WERE NOT ARCHIVED

18 O: WELL, I AM REALLY NOT THAT SURE OF HOW THE ARCHIVEING OF
19 THE FILES ARE DONE SO I WOULD SUGGEST THAT YOU SEND A MESSAGE TO
20 [name2] IN REGARDS TO THIS PROBLEM. THE ARCHIVEING IS DONE
21 DURING THE SWING SHIFT SO IN ORDER THAT YOU DO DON'T NOT GET
22 MISLED BY SOMETHING I AM NOT CERTAIN ON I WOULD RATHER
23 YOU DISCUSS THIS WITH [name2]. HE IS OUT OF HIS OFFICE AT THIS
24 TIME BUT YOU COULD SEND HIM A MESSAGE AND I AM SURE HE WILL GIVE
25 YOU ALL THE INFORMATION YOU WOULD LIKE. [operator's name]

26 U: OK [operator's name]. CAN YOU PEOPLE GO INTO ANY DIRECTORY AND DELETE FILES WHEN YOU
27 NOTICE THAT THEY ARE NOT BEING USED.. [user's name]

28 O: I AM REALLY NOT SURE, HOWEVER I KNOW THAT THAT WOULD NOT BE
29 RESPONSIBILITY WE WOULD TAKE UPON OURSELVES TO JUDGE WHETHER OR
30 NOT YOU WANT YOUR FILES OR NOT.

31 [outside interruption here]
32 O: [user's name], ARE YOU THERE?

33 LINK FROM [user's name], JOB 21, TTY 22

34 U: SORRY [operator's name]
35 WE GOT DISCONNECTED....CAN YOU RECOVER THOSE FILES FOR ME..
36 AS FAR AS I KNOW THEY WERE IN THE DIRECTORY ON THE 16TH..THE NAMES
37 ARE. ...[file name1].DAT;1ABDLLLL CHECK THAT ...[file name1].DAT;1,2
38 [file name2].F4;1....AND [file name3].F4;1.....[user's name]

39 O: OK HOLD ON JUST A MINUTE AND I WILL TRY TO FIND THEM

40 U: RIGHT

41 [outside interruption here]

42 O: ARE YOU STILL THERE?

43 U: RIGHT [operator's name]

44 O: OK YES I HAVE FOUND THE FILES YOU ARE CONCERNED ABOUT, ON THAT
45 FIRST ONE I ASSUME THE ABDLLLL WAS AN ERROR RIGHT?
46 U: RIGHT THAT SHOULD HAVE BEEN [file name1].DAT;1 AND 2.

47 O: OK THEY ARE HERE

48 [outside interruption here]

49 O: ARE YOU THERE?

50 U: RIGHT

51 O: OK I HAVE FOUND THE FILES YOU WANT

52 [outside interruption here]

53 O: OK I HAVE FOUND THE FILES YOU WANT I WILL RETRIEVE THOSE FOR
54 YOU ALSO, OK

55 U: GREAT..WAS THAT [name3] YOU WERE TALKING ABOUT?

56 O: NO, IT IS [name4], HOWEVER WHEN YOU SEND MESSAGE

57 SEND TO [name2]

58 U: RIGHT...IS [name5] STILL AR[outside interruption here]

59 IS [name5] STILL AROUND.

60 O: NO HE ISN'T. HE'S BEEN GONE FOR ABOUT THREE WEEKS NOW.

61 U: OK, THANKS FOR YOUR HELP , [operator's name] HAVE A GOOD DAY [user's name] OUT.

62 O: OK YOU DO THE SAME WILL DO THOSE FILES RIGHT AWAY. BYE

63 U: THANKS AGAIN

64 BREAK

REFERENCES

Mann, W. C. Man-Machine Communication Process. Research. (a proposal from) USC/Information Sciences Institute to ARPA HRRO, October 1974.

Minsky, M. A Framework for Representing Knowledge. M.I.T. AI Memo 306, 1974.

Rumelhart, D. E. Notes on a Schema for Stories. In D. G. Bobrow, & A. M. Collins (Eds.), *Representation and Understanding*, 1975.

Schank, R. C., & Abelson, R. P. Scripts, Plans, and Knowledge. Submitted to *International Joint Conference on Artificial Intelligence*, 1975.