DOCUMENT RESUME

ED 155 622 CS 004 139

AUTHOR	Frederiksen, Carl H.
TITLE	Discourse Comprehension and Early Reading.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Apr 76
CONTRACT	400-75-0049
NOTE	80p.; Paper presented at the Conference on Theory and Practice of Beginning Reading Instruction, University of Pittsburgh, Learning Research and Development Center, April 1976; Not available in hard copy due to marginal legibility of original document; For related documents, see CS 004 132-133, CS 004 135, CS 004 137-173, ED 125 315 and ED 145 399
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	*Beginning Reading; *Cognitive Processes; *Comprehension Development; *Connected Discourse; Decoding (Reading); Discourse Analysis; Primary Education; *Reading Comprehension; *Reading Instruction; Written Language
IDENTIFIERS	Oral Language

ABSTRACT

The primary goal of early reading instruction, according to this paper, should be to teach children to comprehend written discourse in a manner similar to that for oral discourse because both types of discourse require decoding ability--graphic or acoustic. The paper asserts that to simply design reading instruction to achieve the subgoal of decoding (which seems most often to be the case) is likely to lead to subsequent difficulty in achieving the primary goal of comprehension. The paper describes a five-stage system of discourse comprehension, ascending from the basic level of graphic input to the highest level where inferred propositions (comprehension) are processed. Illustrations of these text based inferences as they are used by children in telling and retelling stories show how the systematic study of discourse comprehension can be applied to beginning reading instruction. (The discussion following presentation of the paper is attached.) (RL)

Discourse Comprehension and Early Reading

Carl H. Frederiksen

National Institute of Education

1200 19th Street N. W.

Washington, D. C. 20208

Conferences supported by a grant to the Learning Research and Development Center from the National Institute of Education (NIE), United States Department of Health, Education, and Welfare, as part of NIE's Compensatory Education Study. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred. NIE Contract #400-75-0049

This paper was presented at the conference on Theory and Practice of Beginning Reading Instruction, University of Pittsburgh, Learning Research and Development Center, April 1976.

DISCOURSE COMPREHENSION AND EARLY READING[1]

Carl H. Frederiksen

National Institute of Education
1200 19th Street N. W.
Washington, D. C. 20208

Comprehension and Early Reading

There are two points of departure in approaching beginning reading instruction. One could, on the one hand, start with an understanding of what a child's capabilities are in producing and comprehending oral language; or one could begin with a conception of the ability one wants the child to acquire—proficiency in comprehending written discourse—and attempt to design reading instruction to promote the development of this ability. Ideally, one would approach reading instruction with an understanding both of the child's current linguistic ability, especially the ability to comprehend oral discourse, and of the process of skilled reading. Yet, we still know very little about the cognitive processes underlying the ability to understand discourse, either in children about to begin reading instruction or in skilled readers. Thus, beginning reading instruction has in the past been based more on untested assumptions about the nature of the comprehension process and what is required to learn to comprehend written text, than on an adequate understanding of how children comprehend oral discourse or learn to comprehend written discourse.

The present chapter has two principal objectives. First, after briefly characterizing present theories and knowledge about discourse comprehension, it will attempt to identify goals for beginning reading instruction which take these theories and knowledge into account. In particular, it will establish that the primary goal of early reading instruction should be to teach children to comprehend written discourse in a manner similar to oral discourse, the principal difference being that reading involves decoding graphic information while oral language comprehension involves decoding acoustic information. A necessary sub-goal, then, is to teach graphic decoding. However, it will be argued, designing instruction in early reading solely to achieve the sub-goal of graphic decoding is likely to lead to subsequent difficulty in achieving the primary goal of comprehension, especially in children who have difficulty in learning to decode.

Second, certain recent theoretical and methodological developments in the study of children's discourse comprehension will be described which can contribute to our knowledge of the process of oral discourse comprehension in children beginning to read and of the relationship of oral discourse processing to the processing of written discourse in early reading. The theory has two components: (1) a representation of the propositional knowledge structures which underlie the production and comprehension of discourse; and (2) a taxonomy of text-based inferences in discourse comprehension. The taxonomy identifies classes of inferential operations which may be applied to propositions given in a discourse to generate new (inferred) propositional structures. Since the first component of the theory has been presented elsewhere (Frederiksen, 1975a), the present chapter will concentrate on the second component--text-based inference.

The methodological development consists of coding procedures which allow one: (1) to determine from a child's story (or other discourse) recall precisely what propositional information a child has recalled from a story, and (2) to investigate the specific inferential operations a child has employed in comprehending and retelling a story. Text-based inferences are investigated by comparing those propositions in a child's story recall which are not explicitly represented in the story presented to the child, to those which are explicitly represented in the story. The method provides a very detailed account of the kinds and amount of semantic information a child acquires from a story, and of the inferences he makes in comprehending the story. Thus, the method provides sensitive indices of how a child processes a story during reading or oral discourse comprehension.

Finally, the paper will briefly consider how these developments may contribute to accomplishing the primary goal of comprehension in early reading. Applications to the assessment and diagnosis of differences between children's processing of oral and written discourse and to beginning reading instruction will be discussed.

Comprehension and Early Reading

The conventional way to think about the reading process (and the process of oral language comprehension) is to begin by distinguishing different processing levels which are associated with different internal representations of graphic (or acoustic) input information. All that is required to establish distinct processing levels is to establish that a linguistic input is represented by means of different abstract internal codes such that each code is distinct from the others and the codes may be generated from one another in sequence. Figure 1 summarizes five processing levels associated with four distinct internal codes. Since there is substantial evidence for the existence of each of these internal codes and since each code may be derived from a code earlier in the sequence, the processing levels in Fig. 1 may be regarded as given and taken as a point of departure in thinking about the comprehension process.

Insert Fig. 1 about here

The first level of internal code represents a graphic (or acoustic) input as a set of abstract visual (or acoustic) features. The existence of an abstract feature code has been established by research on visual and auditory information processing (cf. Lindsay & Norman, 1972) and by many experimental studies of (e.g.) letter and word recognition (cf. Neisser, 1957; Gibson and Levin, 1975).

The neural processes which operate on a graphic (or acoustic) input and produce an encoding of the input information as a set of abstract visual (or auditory) features are referred to collectively as <u>feature analysis</u> in Fig. 1.

The second processing level involves the generation of an internal information structure consisting of a sequence of linguistic units--normally a string of lexical concepts and grammatical morphemes--from the available feature information. The process which generates this abstract <u>language code</u> from a set of graphic (or acoustic) features will be referred to as <u>graphic</u> (or <u>acoustic</u>) <u>decoding</u>. Decoding is a complex pattern recognition process which operates both on the abstract feature code and on an <u>internal lexicon</u> which contains information about words--their pronunciation, orthographic structure, syntactic categories, meaning, and pragmatic information about word usage. Decoding itself may involve additional intermediate internal codes consisting of sequences of units such as phonemes (phonemic code) or orthographic patterns (orthographic code). By far the most attention has been given to decoding processes in research and instructional design related to early reading (cf. Gibson & Levin, 1975). Research issues concerning the nature of the decoding component of discourse comprehension will not be the concern of the present paper. No doubt they will be dicussed by others at this conference (c.f., Venesky and Massaro's chapter in the present volume).

The third level involves the generation of an internal representation of the syntactic structure of a language string. While psycholinguistic research has established beyond a doubt that such a representation occurs, the specific nature of the syntactic representation and of the syntactic component of the comprehension process is an open question. While many psycholinguists in the past have accepted the representation of sentence structure associated with generative transformational grammar as <u>the</u> internal representation of sentence structure, the notion that a grammar ought to be "psychologically realistic" has now been accepted by linguists and psycholinguists associated with the generative grammar tradition, leading them to consider alternatives to generative grammar which appear to be more plausible psychologically.[2] Computational approaches to syntax appear to be particularly promising as psycholinguistic models of syntactic processing (e.g., Woods' (1970) and Kaplan's (1972) augmented transition network grammars), especially if they are designed to operate as components of systems which generate semantic interpretations of sentences (e.g., Winograd's (1972) procedural grammar).

The fourth and fifth levels of processing involve the representation of the semantic (propositional) "content" of a linguistic input. The first of these, the <u>interpretive component</u>, takes the syntactic structure of a language string and uses it to generate a <u>propositional structure</u> consisting of a network of concepts and semantic relations linking these concepts into semantic networks. The attempt to specify semantic structures for English sentences is a problem that has occupied investigators in linguistics (e.g., Chafe, 1970; Leech, 1969; Grimes, 1975), computational linguists (e.g., Simmons, 1973) and artificial intelligence (e.g., Schank, 1973; Winograd, 1972). Psychologists have approached the problem as one of specifying the form in which propositional information is represented in memory (cf. Crothers, 1975; Frederiksen, 1975a; Kintsch, 1975; Norman & Rumelhart, 1975). The most interesting attempts to specify models of the process of semantic interpretation have been made by computer scientists who have attempted to program computers to answer questions or in other ways demonstrate an ability to "understand" English sentences (cf. Winograd, 1972; Schank, 1973).

The fifth and highest level processing component operates entirely on propositions, generating new propositions from propositions which are given, e.g., from prior discourse, from discourse context, or from previously acquired knowledge about the world. Any such process will be refered to here as inference. Note that there is no implied reference to truth-value or conditions of valid inference in the present definition of inference. Any proposition which can be generated from one or more given propositions by means of specified operations (to be discussed later) will be referred to as an inferred proposition. Inferential processes in discourse comprehension have become a central topic among workers in the fields of artificial intellegence and computational linguistics who are interested in building "intelligent" language understanding systems (Schank & Rieger, 1974; Rieger, 1975, 1976; Wilks, 1975; Winograd, 1972; Collins, Warnock, Aiello, and Miller, 1975; Collins, 1976). In addition to providing theories of how inference operates in language comprehension, this work makes clear the extent to which inference is involved in the everday processing of natural language discourse. Furthermore, as was indicated in an earlier paper (Frederiksen, 1976a), propositions which are inferentially related frequently occur in discourse although the inferential relations among propositions may not always be explicitly expressed. The existence of high-level "text macro-structures" involving such inferentially related propositions has been recognized by a number of contributors to the field of discourse analysis (e.g., Crothers, 1975; Rumelhart, 1975; van Dijk, 1976; Kintsch & van Dijk, 1976). This fact together with the fact that a host of language processing problems at the discourse level (e.g., anaphora) require inference for their solution, establish that a discourse comprehension system must embody high-level inferential processes.

While discourse comprehension certainly involves the processing levels indicated in Fig. 1, the most important questions about the nature of discourse processing remain to be answered. What is the form of internal representation or internal code at each processing level? How does each component process operate? How do these component processes interact? This last question is the key to understanding how the comprehension process operates as a whole. It is the complex interactions among component processes that have posed the greatest theoretical challenge to those attempting to model the comprehension process in any detail (e.g., Winograd, 1972).

There are two contrasting conceptions of how different components of the discourse comprehension system interact. In one conception, which we shall refer to as the bottom-up conception of discourse processing, decoding, syntactic processing, and semantic interpretation are conceived of as occuring in sequence. The term "bottom-up" refers to the fact that lower level processes occur prior to (and are independent of) higher level processes. In the bottom-up conception, the processes by which a person understands a discourse are controlled by the textual input, that is, there is a more or less automatic parsing of each sentence in an input text followed by semantic interpretation based on sentence syntax. The bottom-up conception has tended to be predominant in the thinking of psycholinguists and reading researchers. The current emphasis on decoding in early reading makes sense if the comprehension process is essentially bottom-up in its organization. However, recent attempts to program computers to "understand" language, e.g., to answer questions and carry out English dialogue, have established clearly that language comprehension systems must operate in a more complex fashion than is consistent with a purely bottom-up conception of discourse processing. For example, while Winograd's (1972) computer program to understand natural language is organized around a syntactic parser, it requires semantic and deductive (inferential) routines which interact with the parser in complex ways. Thus, although the parser in Winograd's system calls these other components when it is necessary to complete syntactic anlaysis and semantic interpretation (e.g., in determining pronominal reference), his system incorporates the notion that the process of comprehension involves the ability of a language user to combine syntactic, semantic, and inferential procedural knowledge in an <u>interactive</u> fashion to produce a semantic interpretation of sentential inputs.

The second conception represents an opposite extreme in which the syntactic and interpretive components are presumed to be always under the control of high-level inferential processes. For example, in Schank's (1972, 1973) system, comprehension is regarded as a process of mapping from grammatical and lexical information contained in input sentences to "conceptual dependency networks." A conceptual dependency network is generated directly from a (minimal) syntactic parsing of an input sentence and from world knowledge and knowledge about the context of an utterance. Lexical verbs are transformed or decomposed into primitive "case frames" containing slots which can be filled by input lexical concepts or other stored concepts. By identifying a verb, retrieving its primitive case frame and assigning concepts from an input sentence or sentence context into "slots" in the frame, the system operates primarily at an inferential level, using lower level components as necessary to "instantiate" the stored frames, thus building up a conceptual dependency network which "fits" the input "data." Thus, the system operates primarily in a top-down fashion. Two aspects of Schank's system, the notion of a frame structure and the use of inference in semantic interpretation, have proved to be extremely powerful ideas (cf., Min[illegible] Wilks, 1975; Schank & Rieger, 1974).

The psychological evidence for top-down processes in discourse comprehension goes back to Bartlett's (1932) original experiments on constructive memory for text. The notion of a frame structure is not unlike Bartlett's notion of a schema (Bobrow & Norman, 1975). Bransford and Franks (1971) revived Bartlett's "constructive" approach to discourse comprehension with the demonstration that in understanding sentences subjects generate new information which was not explicit in the input sentences. Furthermore, they do not discriminate between the two kinds of information, that which was explicit in the text and that which was not. In our own work (Frederiksen, 1975b) we have investigated inferred propositions which are present in subjects' discourse recalls, establishing that inferences are generated during input processing. Furthermore, we (1975c) have manipulated contextual factors which ought to control the extent of inference and produced the expected contextual effects on extent of inferential processing. It was possible to conclude that most of these inferences were generated during input processing. In another study (Frederiksen, 1976a) we have established that propositions presented early in a text affect the processing of propositions presented later in the text to which they are inferentially related.

The research which has been described is consistent with the following assumption:

> Assumption 1. A normal characteristic of skilled reading and oral discourse comprehension is the occurrence of high-level inferential processes which interact with other component processes in a top-down manner.

A principal reason why an efficient discourse processing system has to have top-down characteristics is that such processing is essential to eliminate the enormous processing load which discourse would place on a bottom-up system. The adoption of top-down processing strategies is a powerful means of reducing such intolerable processing load. Examples of top-down strategies are the occurence of text-based inferences in discourse comprehension, inferential control of semantic interpretation and syntactic processing, and the use of contextual information to facilitate decoding.

Assumption 1 naturally leads one to ask if there are conditions which would lead a person to process linguistic inputs in a predominately bottom-up manner. While many exponents of the top-down conception would argue that language processing is always top-down, there may be conditions which would lead to a more bottom-up mode of process interaction in discourse comprehension. Related to the present question is the hypothesis which has been proposed recently (LaBerge and Samuels, 1974; Perfetti and Lesgold, 1975) that persons who are not proficient in decoding will fail to comprehend because, after they have finished decoding a message, they will have inadequate processing resources remaining to complete higher levels of syntactic and semantic analysis. While this hypothesis reflects a bottom-up conception of discourse processing, a second assumption may be offered which proposes a more profound effect of encountering difficulty in decoding or other low-level processing on the discourse processing system:

Assumption 2. If a person encounters difficulty in decoding or other low-level discourse processing, he is likely to revert to a bottom-up mode of discourse processing with respect to all remaining processing levels, thus failing to employ those inferential and other top-down processing strategies which are characteristic of his normal discourse processing.

Assumption 2 predicts that encountering difficulty in decoding will affect not only what processing resources remain, but also the nature of the comprehension process itself. Similarly, it predicts that encountering difficulty in processing sentence syntax or with semantic interpretation due, e.g., to unfamiliar content would have similar effects. If this assumption is valid, then we would expect skilled and unskilled readers to differ not only with respect to decoding proficiency; we would also expect that unskilled readers would learn to process written discourse in such a way that they would fail to apply the powerful top-down inferential processes they have available for processing oral discourse to the processing of written text.

While there is as yet very little evidence bearing directly on Assumption 2, there is a body of research which is suggestive and consistent with this assumption (cf. Gibson & Levin, 1975). Gibson and Levin conclude their review of this literature with the following summary:

> "Good and poor readers do not necessarily differ in the ability to transform a written word to speech, in other words, in the mechanics of reading. The ability to use larger units and to make inferences from the text is involved in skilled reading rather than simply the ability to decode, so that skilled decoders who have not learned to organize the text into higher-order groupings may still be poor readers, so far as comprehension is concerned" (Gibson & Levin, 1975, p. 391).

What Assumption 2 offers is an explanation of why good and poor readers who differ only minimally in decoding ability would be expected to exhibit such striking differences in the ability to comprehend written text.

What, then, is the nature of reading acquisition during the early grades? According to the present view, learning to read involves both the acquisition of graphic decoding processes and learning to apply the full power of the oral discourse processing system to written text. The problem is that even if beginning reading instruction were able to make a child as efficient in graphic decoding as in acoustic decoding, this would be no guarantee that the child would process written text as efficiently or in the same manner as oral discourse. In fact, instruction designed to efficiently teach decoding subskills may actually have demand characteristics which cause children to approach reading in a relatively bottom-up manner whether or not they experience difficulty in decoding. Even the seemingly inocuous task of oral reading may bias a child to approach reading as a task essentially different from that of oral language comprehension. The argument for teaching decoding directly is not that a child cannot otherwise learn to decode (cf. Söderbergh's (1976) case studies of preschool children learning to decode spontaneously); it is that there is an increase in efficiency if decoding is taught directly, the implication being, consistent with a bottom-up conception of discourse processing, that making the child an efficient decoder will result automatically in the child's being able to apply his oral language skills to written text. However, if the increase in efficiency in teaching decoding is bought at the expense of producing a child who processes discourse in a manner which is fundamentally different from how he processes oral discourse, it is not worth the cost.

Early reading instruction, then, must be oriented towards achieving the following two goals:

(1) the primary goal of early reading instruction is to teach children to process written discourse in the same manner as oral discourse; and (2) a subsidiary goal is to assist the child in developing efficient processes for decoding written language. The problem is to design beginning reading instruction to simultaneously accomplish both goals. However, if attainment of the primary goal necessitates some inefficiency with respect to the secondary goal, then that inefficiency must be tolerated.

If beginning reading instruction is to accomplish both of these goals, it must be based on an understanding of the processes children have available to them for comprehending oral discourse as they begin to read, of developmental changes which take place in oral discourse processing during the period of early reading, and of changes which take place in the processing of written discourse as a child learns to read. It will also require that we have available procedures for determining how a child is processing oral and written discourse and whether he is processing both oral and written discourse in the same manner. In the next section, some recent theoretical and methodological developments will be presented which provide both an approach to describing children's processing of oral and written discourse and a procedure which is adaptable for use in the classroom for determining how a child differs in his processing of oral and written discourse. The story retelling task which is employed is both natural for children and suitable for use as an instructional task.

Text-Based Inference in Children's Story Comprehension

The previous discussion identified five processing levels in discourse comprehension on the basis of different internal representations of a linguistic input. Comprehension was described as a process by which a semantic (propositional) representation is generated from input discourse. Acquired propositional knowledge was described as falling into two categories—that which results directly from semantic interpretation of an input text and that which is not explicitly represented in a text and hence must be inferred by a listener or reader. To describe the processes of semantic interpretation and inference in any detail will require that we first be able to specify the explicit propositional content of a discourse.

A natural approach to the problem of specifying the propositional content of discourse is to consider the other half of the communication process—discourse production. Figure 2 presents a conception of the process of discourse production in which a text is viewed as resulting from a series of communicative decisions whereby a speaker (or writer) generates discourse from his store of "message relevant" conceptual and propositional knowledge.[3] The store of semantic knowledge from which a textual message is derived will be referred to as a message domain. In the present conception, discourse production involves three levels of decision, each of which determines a different aspect of discourse structure—semantic (propositional) content, textual organization and cohesion, and sentence structure. At the first decision level, a speaker (or writer) retrieves units of semantic information for explicit incorporation into a discourse. This selection process reflects pragmatic assumptions a speaker (or writer) makes about what knowledge he already shares with the person being addressed and about the inferential capabilities of the addressee. That propositional knowledge which a speaker (or writer) selects for explicit incorporation into a message is called a message base. At the second level of decision, a speaker (or writer) makes staging decisions which determine how the selected units of semantic information are to be organized into discourse, including decisions about sequence, topicalization, reference, and correspondence between semantic units (such as propositions) and textual units (such as sentences)[4]. The resulting "staged" message base is referred to here as a text base, indicating that it contains both "textual" and semantic information. Finally, a speaker applies his knowledge of sentence structure to generate a sequence of sentences from the text base. Each of these decision levels involves communicative decisions which successively reduce the amount of free variation in text. Presumably, most free variation is eliminated at the level of the text base; the last stage, text generation, involves only the application of grammatical rules which are appropriate in a speaker's (or writer's) language community. Of course, discourse production occurs in real time with many interactions among component processes (cf. Frederiksen, 1976a).

Insert Fig. 2 about here

To illustrate, Figure 3 contains an example of a message base for the following children's story (from Hall, Reder, & Cole, 1975):

The Flower Pot Story

(1) This is Michele. She is watering the flowers.

(2) Crash! Now Michele thinks that Mother will be mad. She wants to run away.

(3) "I'm not mad," says Mother. "I know you didn't mean to do it. Let's clean up the mess."

(4) Michele picks up the flowers. She gives them to Mother. "Don't worry," says Mother, "we'll put them in a nice pot."

(5) Now the flowers are okay and the mess is all cleaned up. "Come on," says Mother. "Let's go and make some cookies."[5]

Each numbered row in Figure 3 denotes a proposition consisting of a network of <u>concepts</u> (in parentheses) connected by labeled <u>semantic relations</u> (the arrows connecting the concepts); each proposition represents an <u>event</u> or <u>state</u>. For example, proposition 100 represents the event <u>Michele</u> caused the action <u>water</u> to affect the object set <u>flowers</u>. The action in proposition 100 was initiated in the past with continuous aspect as indicated by the operators TEM and ASPCT (cont) on the ACT relation (each of which is denoted by an @ sign followed by an operator label). A concept may be an object (in which case it is preceeded by a colon in the Fig.). Objects appear both within propositions and on an object list together with relations that indicate how each object is determined and quantified. For example, Michele is definite singular. Every other concept is an action, attribute, degree, location, or time. Propositions frequently are embedded within other propositions (as indicated by a square bracket in a proposition containing a proposition number). For example, in proposition 203, Michele is the patient of the processive action want which has as a goal the event represented by the embedded proposition 204. Full details of these network structures for representing propositional knowledge are given in Frederiksen (1975a).

- - - - - - - - - - - - -
Insert Fig. 3 about here
- - - - - - - - - - - - -

Staging decisions explain why, e.g., sentences (1) were generated from proposition 100 rather than some other sentence or sentences. This particular staging establishes Michele as the main topic of the story and the action as subsidiary comment. An alternative staging might have been Michele is watering the flowers (correspondence between the textual unit sentence and the semantic unit proposition) or Here are the flowers. They are being watered by Michele (topicalization of the flowers). A detailed theory of staging remains to be worked out but the outlines of a theory have been given by Grimes (1975) and summarized by Clements (1976). Presumably, once staging decisions have been made, a text may be generated by applying grammatical rules which map from propositions + staging information (the text base) to English sentences. Note, that in research on discourse comprehension, it has been common to confuse aspects of staging with aspects of the propositional message base.

Suppose that the Flower Pot Story was read to a child (or alternatively, the child read the story) and the child was then asked to retell the story. The child's recall was recorded and transcribed. For example, the following recall was obtained from a four-year old in the study by Hall, Reder, and Cole (Subject 1):

> A little girl was watering the flowers and then she was ...that she was watering the flowers. She ...she broke the glass. Don't worry said Mom, we'll clean up the mess ....she won't get mad. She put it in a nice new pot and then she was cleaning up the pot, and that was only a accident also. The flowers are okay, and the mess is all cleaned up.

We want to determine first what information from the propositional structure in Figure 3 was recalled by this child. The procedure which we use is to code the child's recall against the propositional data structure, checking off every concept or relation which was recalled. Thus, we obtain a detailed account of what semantic information the child has recalled from the story. Figure 4 presents a list of propositions from Figure 3 which were recalled by this child. Notice that in addition to recalling a portion of the story content, this child has included propositions which are related to but are not identical to propositions in the message base. The attempt to analyze propositions such as these leads directly to a theory of text-based inference in discourse comprehension.

Insert Fig. 4 about here

Before describing our approach to text-based inference, it will be helpful to characterize further the propositional structures on which inferences are based.[6] A propositional structure consists of a set of concepts connected into networks by labeled binary semantic relations. A relation is defined in terms of a triple consisting of a pair of concept slots and a connecting relation. For example, the relation ()--CAT-> () connects two object categories such that the object category in the right slot is a subset of the object category in the left slot, e.g., (:BIRDS)--CAT-> (:CANARIES). The two smallest semantic units are thus: (1) lexical concepts and (2) relational triples consisting of pairs of concepts connected by semantic relations. All higher-order units of semantic information are composed of relational triples.

Higher-order units occur at several levels of rank or complexity and will be described in their order of complexity (cf. Frederiksen, 1976a).

The next largest semantic unit, (3) an *event frame*, is analogous to a case frame in Schank's (1972) theory and is composed of a system of relational triples which are connected to an action and identify the various participants in the action, its resulting effects, etc. An example of a resultive event frame is found in proposition 100. Here, the event frame consists of everything except the time (tense and aspect) information. Note, that this is an instance of an incomplete event frame; it represents only that part of the frame which was explicitly expressed in the story. The full event frame is

```
(:MICHELE)--AGT-> ('WATER)*┬-OBJ1-> (:FLOWERS)
                           ├-SOURCE->["101]
                           └-RESULT->["102]
```

where embedded propositions 101 and 102 represent the unspecified state of the flowers prior to the action *water* and after the action *water* has taken place, respectively. Proposition 500 contains an example of an event frame with an unfilled slot—the agent slot. One kind of inferential operation might be to fill this slot with an animate object. Proposition 201 contains an example of a processive event frame:

```
(:MICHELE)--PAT-> ('THINK)--THEME2-> ["202]
```

where 202 is an embedded proposition specifying the content or "theme" of the cognitive process think.

The next largest semantic unit, (4) a proposition, represents an event or state. An event proposition is composed of an event frame together with additional relations which further identify the event, e.g., by specifying the time and location at which the event took place, or by further specifying the nature of the action. Examples of stative propositions in the Flower Pot Story are propositions 205 (locative) and 408 (attributive). Other stative propositions are 206, 407, 501, and 505. All other propositions in the story represent events. Proposition 501 is an example of an event proposition consisting of an event frame containing embedded propositions, a tense operator, and a locative relational triple (a / indicates a branch in the network at the point indicated by an *).

The two units which are highest on the rank scale are composed of propositions and connect propositions either with algebraic relations or with dependency relations, that is, logical, causal, or conditional relations. A unit composed of a pair of propositions connected by an algebraic relation is called (5) a relative system. A relative system may specify relative time, location, or comparative information involving attributes of objects or actions. An example of a relative system, in this instance one involving relative location, is given by proposition 207 which specifies that the location of Michele after the action run (206) is not the same as her location prior to the action (205). (Here, a # sign is used to indicate a label for an unfilled slot.) The algebraic relation proximity connects slots in 205 and 206 and has the negative operator NEG applied to it.

(6) a dependency system consists of a pair of propositions which are connected by means of logical, causal, or conditional relations--relations which indicate dependencies among propositions. Dependency systems are of three types: logical, causal and conditional systems. No examples of dependency systems occur in the Flower Pot Story, but examples may be found in Frederiksen, 1975a.

Now consider the first sentence in the story recall of subject 1:

> A little girl was watering the flowers and then she was ...that she was watering the flowers.

What information has this child recalled from the story? And, how has she operated on the propositions in the story to generate this sentence? Figure 4 lists the propositions from Fig. 3 which are represented in this child's story recall. The above sentence reproduces the information from proposition 100 which is indicated in Fig. 4, and also includes two alterations of the proposition. First, the object class occurring in the AGT slot, Michele, has been replaced with a little girl, a less specific class of children.

That is, Michele refers to a definite little girl while a little girl refers to any token from the class of little girls. Thus, the object class in the AGT slot in the child's proposition includes the object class in proposition 100 as a subset, and the child's proposition is thus more general than that given in the story—it includes proposition 100 as a special case. Second, the child has changed the time reference from present to past. Both of these changes which occur in the child's text are evidence for inference—processes which operate on given propositions to produce new propositions. The occurence of such operations in a child's story recall is evidence for discourse processing at the inferential level. The kind and amount of such text-based inference ought to be indicative of the nature of discourse processing in these children, large amounts of inference being evidence for top-down processing strategies. The detailed description of children's text-based inferences and of the semantic information they acquire from discourse ought to contribute significantly to describing the processing of both oral and written discourse during the period of early reading and developmental changes in both processes.

The classification of inferences such as these involves examining the relationships between propositions in a child's story recall and propositions in the message base for the story presented to the child. A child's propositions may be classified on the basis of the particular inferential operations which can be applied to propositions in the message base to generate the child's propositions. Types of inference correspond to classes of operations on propositions. Since in some instances it may be possible to generate a child's propositions in more than one way from propositions in a message base, the convention will be adopted that in classifying a child's inferred propositions, inferential operations are applied to those propositions in the message base which are most closely related to the child's propositions (that is, the fewest possible operations are applied in classifying each of a child's inferred propositions).

Table I presents a summary of the inference types which we have identified thus far. This classification is intended to be exhaustive and is based on considerations of what operations are possible given the nature of propositional structures and on analyses of inferred propositions in children's story recalls. Eight major classes of operations on propositions have been identified which operate on different semantic units and/or involve different operations on these units. Major classes of operations are subdivided into more specific categories of inferential operations. The resulting classification consists of twenty-six inference types (see Table 1). In the full classification of text-based inferences, the inference types themselves are further subdivided (cf. Frederiksen, 1976b). The classification will be illustrated by means of examples in what follows.

Insert Table 1 about here

The first major class of inferences, lexical operations, operates at the conceptual level. Two types of operations can occur: (1) lexical expansion, expanding a lexical concept into one or more propositions; and (2) lexicalization, replacing one or more propositions with a lexical concept. For example, if in retelling The Flower Pot Story, a child said (a) Let's clean up the dirt and pieces of flower pot all over the floor instead of (b) Let's clean up the mess, the child would have expanded the lexical concept mess. If the child generated (b) from (a), he should have lexicalized the propositions underlying (a).

The second major class of inferences, identification operations, involves operation on objects, actions, stative propositions, or events which further specify or identify an object, action, state, or event. For example, identifying an object involves providing stative information about the object which distinguishes it from other objects. Six types of identification operations may be distinguished:

(3) attribute inference: specifying an attribute of an object or action, e.g.,

stative attribution: Story 1? We'll put them in a nice pot. (proposition 408); Protocol (subject 1): She put it in a nice new pot. (pot is identified by the attribute new)

manner attribution: Story 2: Now Jimmy's mad ...That makes him feel better. (proposition 520); Protocol (subject 5) Jimmy was happy. (the attribute happy is attributed to the processive action feel in 520);

(4) category inference: classifying an object or action into a category, e.g.,

Story 2: He is buying an ice cream. (ice cream); Protocol (subject 1): he was buying a popsicles of ice cream. (popsicles are a subset of the category ice cream);

(5) time inference: specifying a time or duration for an event or state, e.g.,

(a) PRES -> PAST:

Story 1: She is watering the flowers. (proposition 100)

Protocol (subject 1): A little girl was watering the flowers;

(b) TENSELESS -> FUT:

Story 1: Let's clean up the mess. (proposition 306);

Protocol (subject 1): we'll clean up the mess;

(c) ASPCT: COMPLETIVE -> CONT:

Story 1: The mess is all cleaned up. (proposition

500);

Protocol (subject 1): She was cleaning up the pot;

(6) locative inference: specifying a location for an event or state, e.g.,

Story 1: She (Michele) is watering the flowers. (proposition 100); Protocol: Michele is watering the flowers on the windowsill;

(7) part-structure (HASP) inference: specifying a part of an object, e.g., the flower's blossoms;

(8) degree inference: specifying a degree of an attribute, e.g.,

Story 1: We'll put them in a nice pot. (proposition 408); Protocol: We'll put them in a very nice pot.

The third major class of text-based inferences, frame operations, consists of operations on event frames. Ten distinct categories of operations on event frames can be identified which correspond to inference types (9) through (18) in Table 1:

(9) act inference: filling an unfilled action slot in an event frame, e.g.,

Story 1: Mother will be mad. (proposition 202); Protocol Mother will feel mad. (the cognitive action feel is specified);

(10) case inference: filling a concept slot in an event frame, specifically inserting a concept into an agent, instrument,

dative, or object slot in a resultive event frame; or into a patient, dative, or object slot in a processive event frame, e.g.,

Story 1: Now the mess is all cleaned up (proposition 500);

Protocol (subject 1): She (Mother) was cleaning up the pot.. (empty AGT slot filled with Mother);

(11) instrumental inference: generating a cause of a proposition marked as a result, e.g., if Story 1 said: Mother got mad (proposition 202 marked as a result), and a child's recall said Michele made her mother mad, the child would have supplied an agent and an action for proposition 202;

(12) result inference: generating a proposition indicating the result of an action, e.g.,

Story 2: Take half of mine (ice cream). (proposition 511):

```
(:JIMMY) --AGT -> ('TAKE)┬-OBJ1-> (:HALF.ICE.CREAM.B)
                         └-RESULT -> [ ];
```

Protocol (subject 5): Jimmy, you can have half of mine; fills the result slot with:

(:JIMMY)--PAT-> ('HAVE)--OBJ2-> (:HALF.ICE.CREAM.B);

(13) source inference: generating a proposition indicating a state existing prior to an action, e.g., in the preceding example, a child might say Jimmy you don't have any ice cream. Take half of mine;

(14) goal inference: generating a goal for an action, e.g.,

Story 1: Michele picks up the flowers (proposition 400);

Protocol: Michele picks up the flowers so her mother won't be mad. (specifies a goal for the action pick up);

(15) theme inference: generating a theme for a cognitive action, e.g.,

Story 1: Let's clean up the mess. (proposition 305), Michele picks up the flowers (proposition 400); Protocol (subject 3): Let's pick up the flowers. (proposition 400 is inserted into the theme slot of 305);

(16) frame transformation: transforming a frame of one type into a frame of another type (see Frederiksen, 1976b for examples);

(17) disembedding operations: removing a proposition from an event frame in which it is embedded, e.g.,

Story 1: She (Michele) wants to run away. (proposition 204 embedded in the goal slot of 203); Protocol (subject 4): Michele....did run off. (proposition 204 has been removed from the goal slot of proposition 203);

(18) embedding operations: inserting a proposition into a slot in an event frame, e.g.,

Story 2: This is Jimmy. He is buying ice cream. (proposition 100); Protocol (subject 2): Jimmy wanted to buy ice cream. (proposition 100 has been embedded in the goal slot of a generated event, Jimmy wanted).

The fourth major class of inferences, event generation, involves generating an event frame into which an object or proposition is inserted. Ten sub-types of (19), event inference, occur corresponding to the different slots ("cases") which occur in processive and resultive event frames. Examples of several types of event inference are as follows:

Processive events:

19A. PAT (story 1, subject 1):

Protocol: She (Michele) broke the glass. (Michele is inserted into the PAT slot of a generated event);

19B. OBJ2 (story 2, subject 5):

Protocol: (see example 12A; the generated result contains the concept (:HALF.ICE.CREAM.B) in the OBJ2 slot);

19AB. PAT + OBJ2 (story 1, subject 3):

Protocol: Said mother, come, lets have some cookies. (the generated event contains Mother and Michele in the PAT slot and cookies in the OBJ2 slot);

Resultive events:

19E. AGT (story 1, subject 3):

Protocol: Michele broke (the flowers) (an event is generated having an animate object from the text as AGT—Michele);

19G. OBJ1 (story 2, subject 1):

Protocol: It (the ice cream) got all broken. (the object ice cream is inserted into the OBJ1 slot of a generated event);

19EG. AGT + OBJ1 (story 2, subject 2):

Protocol: he (Jimmy) dropped it (the ice cream).

The fifth major class of text-based inferences, macrostructure operations, involves operations on propositions in a text which result in new propositions which are either more general than (i.e. are superordinate to) or more specific than (i.e. are subordinate to) the propositions in the message base for the text. (20) Superordinate inference involves replacing a concept occupying a slot in a proposition with a superordinate class of concepts (that is, a class of concepts which includes the concept in the proposition as a subset). Numerous sub-types of superordinate inference can be distinguished on the basis of the type of slot and/or concept operated on. (21) Subordinate inference involves the same operations in reverse; a conceptual class is replaced by a concept which is subordinate. Again, there are as many sub-types of operations as there are different slots on which to operate. Examples of superordinate inference are:

B. AGT (story 1, subject 1) - superordinate object class in AGT slot, Story 1: This is Michele. She is watering the flowers; Protocol: A little girl was watering the flowers.

B. OBJ1 (story 1, subject 5) - superordinate object class in

OBJ1 slot, Story 1: She is watering the flowers. (proposition 100); Protocol: She was waterin' the plants.

B. PAT (story 2, subject 1) - superordinate object class in PAT slot, Story 2: (You) come on and fight. (proposition 420); Protocol: let's come on and fight. (PAT slot contains both Jimmy and the other boy).

Examples of subordinate inference are:

A. Subordinate action (story 2, subject 5, proposition 410) Story 2: I'm gonna beat you up. Come on and fight: Protocol: Sock (you) in the head. Let's fight. (the action is subordinate to beat up);

B. AGT (story 1, subject 1, propositions 406-8) - subordinate object class in AGT slot, Story 1: we'll put them in a nice pot; Protocol: She (Mother) put it (flowers) in a nice new pot;

B. OBJ1 (story 1, subject 1, proposition 500) - subordinate object class in OBJ1 slot; Story 1: the mess is all cleaned up; Protocol: She was cleaning up the pot.

The sixth major class of inferences, algebraic operations, involve generating a relative (algebraic) system given a set of relative propositions. For example, if a text specified that two children were naughty and, in retelling the story, a child produced a comparative statement about how naughty the children were the child would have made an algebraic inference. Algebraic inference may involve metric propositions which specify metric attributes (i.e. attributes having a degree), or they may involve relative object classes or propositions which are nonmetric. Examples of each sub-type are:

22A. Connecting metric propositions with algebraic relations, e.g., specifying temporal order:

Protocol (story 1, subject 1): A little girl was watering the flowers and then she...(temporal order);

22B. Connecting nonmetric relative object classes or propositions:

Protocol (story 1, subject 2): her (Michele's) mother (kinship relations are examples of nonmetric algebraic relations between relative object classes, that is, object classes which are defined relative to one another)

Dependency operations, the seventh class of inferences, are operations which connect propositions with dependency relations, that is, relations which establish that one proposition is dependent on another proposition—logically, conditionally, or causally (functionally). Three inference types may be identified within this class: (23) causal inference, connecting propositions by means of causal relations, thus generating a causal system; (24) conditional inference, connecting propositions by means of conditional relations, thus generating a conditional system; and (25) logical inference, connecting propositions by means of logical relations, thus generating a logical system. Again, sub-types can be identified for each inference type. Three illustrative examples are:

23. Causal inference:

A. connect unconnected events with causal relations

Protocol: (story 1, subject 3): Michele was watering the flowers and crashed and broke...

(Michele caused proposition 200)

F. inchoative (story 1, subject 1, proposition 202):

Story 1: Michele thinks that mother will be mad;

Protocol: She (Mother) won't get mad;

24. Conditional inference:

A. Enablement inference (story 1, subject 5) - specify antecedent conditions for an event or state

Story 1: Now Michele thinks that Mother will be mad. She wants to run away;

Protocol: She (Michele) thought her mother might be mad (201, 202). Then she would run away (204-207).

Finally, a child may operate on the truth-value of a proposition. Two types of (26) truth-value operations are possible: qualification and negation. Examples of these sub-types are:

26A. Qualification (story 1, subject 5, propositions 201, 202): Story 1: Now Michele thinks that Mother will be mad; Protocol: She thought her right mother be mad;

26B. Negation (story 1, subject 1):

Story 1: Mother will be mad; Protocol: She (Mother) won't get mad.

To illustrate the coding of inference types, the analysis of text-based inferences which occur in the story recall given previously is presented in Table 2. For each line of text from the child's protocol, the Table presents (1) the number(s) of the proposition(s) in the message base most closely matching the proposition produced by the child in that line of text and (2) the code number (from Table 1) of the inference type(s) corresponding to those operations which must be applied to the indicated proposition(s) to derive the child's proposition. This child's protocol is fairly typical of those obtained by Hall, Peder, and Cole for the Flower Pot Story. Inspection of this Table confirms, by example, that inference is heavily involved in the everyday processing of natural discourse.

Insert Table 2 about here

How much inference actually occurs when four year olds comprehend and retell simple narrative stories? What kinds of inferences occur and how frequently? Are there import individual and developmental differences in the amount and kinds of inferences which children produce? How are text-based inferences affected by the characteristics of a text? While we do not yet have answers to all these questions, some preliminary indication of what we can expect to find is provided in Table 3 which reports the distribution of inferences for five of the children studied by Hall, Peder, and Cole who were asked to retell two short orally presented narrative stories.

Insert Table 3 about here

It is apparant from Table 3 that many of the possible inference types occur in this sample of children's story recalls. While we need to obtain recalls for stories unaccompanied by pictures (since inferred information may be derived from the pictures accompanying the story), the experimental conditions used by Hall, Peder, and Cole were sufficiently natural that it is reasonable to suppose that the inferences observed are not unrepresentative of those which typically would occur in the comprehension of short narrative stories. Of the twenty-six inference types, sixteen occurred in these children's recalls and many different sub-types also occurred. The most frequent class of inferential operations was identifying operations (33), followed by event generation (23), macrostructure operations (18), frame operations (14), algebraic operations (11), dependency operations (5), and truth-value operations (2). No lexical operations occurred. Most identifying operations involved time inference, reflecting a strong tendency to shift the time orientation from the present to the past. Event generation is that category of inferences which is most likely to reflect the effects of pictures. Most of the algebraic inferences involved the temporal ordering of events. While dependency operations were infrequent for these stories, there is reason to expect this category of inferences to be greater for types of discourse other than simple narrative stories. Thus, in answer to the first two questions, there is a substantial amount of inference and the inferences are distributed over a wide range of inference types. What about individual differences? If one looks simply at the total number of inferences produced by individual children for each story, there appear to be very substantial individual differences in the amount of text-based inference. However, there is some instability of these differences across stories. If one looks at the patterns of inference types for individual subjects, there is consistency over subjects for some types (e.g., time inference) and inconsistency for others. The results would appear to indicate that a detailed investigation of individual and developmental differences would be extremely fruitful. Finally, there were very substantial story effects, even for two short narrative stories. The strong suggestion is that discourse characteristics will prove to have important effects on text-based inferences. What is needed here is an investigation of specific discourse characteristics that produce particular kinds of text-based inferences.

Applications to Early Reading Instruction

It remains to consider how the developments in the study of children's discourse comprehension which have been reported here can contribute to the design and practice of early reading instruction. There appear to be three principal ways in which this research can benefit the teaching of reading: (1) by establishing goals for early reading instruction, (2) by providing assessment procedures on which to base instructional decisions, and (3) by providing instructional tasks and procedures which teachers can employ in teaching beginning reading.

Goals need to be established for early reading instruction at many levels. We have already examined the highest level goals and seen that the conception of the comprehension process which is being developed by researchers already has a direct bearing on the establishment of high-level goals. However, more detailed and specific goals must also be established in order to accomplish the major goal of comprehension. Specific goals can be established on the basis of research, of the kind reported here, which investigates the extent and types of text-based inferences which are characteristic of children who are highly successful in school tasks, and on the kinds and amount of semantic information they acquire from written texts. Another kind of specific instructional goal is a goal established for an individual child. Such goals could be based on, e.g., an analysis of the kinds of semantic information children acquire from written texts, including information which is inferred; or on comparisons of the kinds of information a child acquires from written and spoken discourse.

The research which has been described can contribute to the development of assessment procedures in two ways. First, the methods which have been described for analyzing children's story recalls could be adapted for use in classrooms as an assessment technique, much in the way analysis of oral reading "miscues" is employed as a classroom technique. For example, the methods could be used to compare the information a child acquires in reading to that which he acquires from a structurally similar text presented orally. Such a comparison can be of value in making instructional decisions for an individual child. Thus, a teacher could establish reading goals for a child which were consistent with that child's comprehension of oral discourse; or a teacher might identify a child's "comprehension problem" as a general language problem rather than a problem specific to reading. Second, the classification of inference types can be employed as a basis for systematically constructing achievement test items to assess aspects of discourse processing.

The story retelling task which we have been employing in our research is desirable not only as a research task, but also as an instructional task. Discourse recall is natural for children, it is appropriate both for reading and oral discourse comprehension, it does not necessarily bias a child to process a text in a particular way, and it provides a rich source of information about how a child processes a text in comprehending it. Furthermore, as research knowledge based on this task accumulates, that knowledge will be directly generalizable to instruction which employs the same or similar tasks. Finally, another way in which reading instruction can be influenced by research on children's discourse comprehension is through enriching teachers' conceptions of the knowledge and skill that is involved in comprehending discourse.

Figure Captions

Fig. 1 - Processing Levels in Discourse Comprehension

Fig. 2 - Communicative Decisions in Discourse Production

Fig. 3 - Propositional Network: The Flower Pot Story (Story 1)

Fig. 4 - Propositional Network: Subject 1

Footnotes

1. This research was supported by grant number GS-4023 from the National Science Foundation to the author while the author was at the University of California, Berkeley. Any conclusions or other views expressed here are the author's own and do not represent the views of the National Institute of Education, the National Science Foundation, or any other part of the federal government. No official support or endorsement by the National Institute of Education is intended or should be inferred.
2. Conference on "New Approaches to a Realistic Model of Language," Massachusetts Institute of Technology, March 9-10, 1976, co-sponsored by M.I.T. and the American Telephone and Telegraph Company.
3. Figure 2 is reproduced from R. Freedle (Ed.), Discourse production and comprehension. Norwood, N.J.: Ablex Publishing Co., 1976.
4. The term staging was suggested by Grimes (1975) who has drawn an analogy between the "staging" of discourse and the "staging" of a theatrical production.
5. In Hall, Reder, and Cole (1975) the Flower Pot Story was accompanied by pictures illustrating the story. Of course, children's story recalls could include semantic information derived from the pictures as well as from the text.

Page 2 - Footnotes (cont'd)

6. The propositional structures are presented in detail in Frederiksen (1975a); the propositional notation and the characterization of these structures in terms of ranked units are given in Frederiksen (1976a).
7. Story 1 is the Flower Pot Story in Hall, Reder, and Cole (1975) Table 1.
8. Story 2 is the Ice Cream Story in Hall, Reder, and Cole (1975), Table 2.

References

Bartlett, F. C. Remembering. Cambridge: The University Press, 1932.

Bransford, J. D. & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Bobrow, D. G. & Norman, D. A. Some principles of memory schemata. In D. Bobrow and A. Collins (Eds.), Representation and understanding. N.Y.: Academic Press, 1975.

Chafe, W. Meaning and the structure of language. Chicago: University of Chicago Press, 1970.

Clements, P. Effects of staging on recall from prose. Ph.D. Dissertation, Cornell University, 1976.

Collins, A. Processes in acquiring knowledge. In R. C. Anderson, R. J. Spiro, and W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum Assoc., 1976.

Collins, A., Warnock, E. H., Aiello, N., & Miller, M. Reasoning from Incomplete Knowledge. In D. Bobrow & A. Collins (Eds.), Representation and Understanding: Studies in Cognitive Science, N.Y.: Academic Press, 1975.

Crothers, E. J. Paragraph structure description. Report No. 40. Boulder, Colorado: University of Colorado, Institute for the Study of Intellectual Behavior, 1975.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 1[illegible], 7, 371-458. (a)

Frederiksen, C. H. Acquisition of semantic information from discourse: effects of repeated exposures. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 158-169. (b)

Frederiksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology, 1975, 7, 139-166. (c)

Frederiksen, C. H. Semantic processing units in understanding text. In R. Freedle (Ed.) Discourse production and comprehension. Norwood, N.J.: Ablex Pub. Corp., in press. (a).

Frederiksen, C. H. Text-based inference in children's discourse comprehension, in preperation. (b)

Gibson, E. J. & Levin, H. The psychology of reading. Cambridge, MA.: M.I.T. Press, 1975.

Grimes, J. E. The thread of discourse. The Hague: Mouton, 1975.

Hall, W., Reder, S. & Cole, M. Story recall in young black and white children: effects of racial group membership, race of experimenter and dialect. Developmental Psychology, 1975, 11, 628-634.

Kaplan, R. M. Augmented transition networks as psychological models of sentence comprehension. Artificial Intelligence, 1972, 3, 77-100.

Kintsch, W. & van Dijk, T. A. Comment on se rappelle et on resume des histoires. Languages, 1975, 40, 18-116.

LaBerge, D. & Samuels, S. J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Leech, G. Towards a semantic description of English. Bloomington, IN: Indiana University Press, 1969.

Lindsay, P. & Norman, D. A. Human information processing. N.Y.: Academic Press, 1972.

Minsky, M. A framework for the representation of knowledge. In P. Winston (Ed.), The psychology of computer vision. New York: McGraw Hill, 1975, in press.

Neisser, U. Cognitive psychology. N.Y.: Appleton-Century-Crofts, 1966.

Perfetti, C. & Lesgold, A. Discourse comprehension and reading skills. In P. Carpenter and M. Just (Eds.), Cognitive processes in comprehension: Twelfth annual Carnegie symposium on cognition, in preparation.

Rieger, C. The commonsense algorithm as a basis for computer models of human memory, inference, belief, and contextual language comprehension, unpublished manuscript, University of Maryland, 1975.

Rieger, C. The algorithmic basis of language-related human intelligence. Georgetown University Round Table on Languages and Linguistics, 1976, in press.

Rumelhart, D. E. Notes on a schema for stories. In D. Bobrow and A. Collins (Eds.), Representation and understanding: studies in cognitive science. New York: Academic Press, 1972.

Schank, R. Identification of conceptualizations underlying natural language. In R. Schank and K. Colby (Eds.), Computer models of thought and language. San Francisco: W. H. Freeman, 1973.

Schank, R. C. & Rieger, C. J. Inference and the computer understanding of natural language. Artificial Intelligence, 1974, 5, 373-412.

Simmons, R. F. Semantic networks: their computation and use for understanding English sentences. In R. Schank and K. Colby (Eds.), Computer models of thought and language. San Francisco: W. H. Freeman, 1973.

Soderbergh, R. Learning to read between two and five: some observations on normal hearing and deaf children. Georgetown University Round Table on Languages and Linguistics, 1976, in press.

van Dijk, T. A. Macro-structure and cognition. In P. Carpenter and M. Just (Eds.), Cognitive processes in comprehension: Twelfth Annual Carnegie symposium on cognition, in preparation.

Wilks, Y. A preferential, pattern-seeking, semantics for natural language. Artificial Intelligence, 1975, 6, 53-74.

Woods, W. A. Transition network grammars for natural language analysis. Communications of the ACM, 1970, 13, 591-606.

Winograd, T. Understanding natural language. Cognitive Psychology, 1972, 3, 1-191.

Table 1

Major Classes of Text-Based Inference

Class of Operations	Units[a]	Inference Types[b]
I. Lexical Operations	lexical concepts	1. LEXICAL EXPANSION
		2. LEXICALIZATION
II. Identification Operations	objects, actions, states, events	3. ATTRIBUTE INFERENCE
		4. CATEGORY INFERENCE
		5. TIME INFERENCE
		6. LOCATIVE INFERENCE
		7. HASP INFERENCE
		8. DEGREE INFERENCE
III. Frame Operations	event frames	9. ACT INFERENCE
		10. CASE INFERENCE
		11. INSTRUMENTAL INFERENCE
		12. RESULT INFERENCE
		13. SOURCE INFERENCE
		14. GOAL INFERENCE
		15. THEME INFERENCE
		16. FRAME TRANSFORMATION
		17. DISEMBEDDING OPERATIONS
		18. EMBEDDING OPERATIONS
IV. Event Generation	event propositions	19. EVENT INFERENCE

Table 1 (cont.)

Class of Operations	Units[b]	Inference Types[c]
V. Macrostructure Operations	propositions	20. SUPERORDINATE INFERENCE
		21. SUBORDINATE INFERENCE
VI. Algebraic Operations	relative systems	22. ALGEBRAIC INFERENCE
		A. METRIC
		B. NONMETRIC
VII. Dependency Operations	dependency systems	23. CAUSAL INFERENCE
		24. CONDITIONAL INFERENCE
		A. ENABLEMENT INFERENCE
		B. PRESUPPOSITION
		C. ANTECEDENT INFERENCE
		25. LOGICAL INFERENCE
		A. DEDUCTIVE INFERENCE
		B. CONDITIONAL PERFECTION
VIII. Truth-Value Operations	propositions	26. TRUTH-VALUE OPERATIONS
		A. QUALIFICATION
		B. NEGATION

Table 2

Classification of Inferences: Subject 1, Story 1

Text from Protocol	Proposition(s)[a]	Inference Types
A little girl[1] was[2] watering the flowers	100	20B (AGT), 5a (PAST)
and then	100	22A ~~(TEM)~~ (ORD)
she broke the glass.	100	19E
Don't worry said Mom,	403	5A (PAST)
we'll clean up the mess	306	5A (FUT)
she won't[1] get[2] mad.	202	26B, 23F
She[1] put[2] it in a nice new[3] pot[4]	406-408, (:POT)	21B (AGT), 5A (PAST), 3A, 17D (THEME)
and then	406, 500	22A (TE[illegible] (ORD)
she[1] was[2] cleaning up the pot[3]	500	10A, 5A (ASPCT-CONT), 21B (OBJ1)
and that was only a accident also.	302, 303	19A

---

[a] Proposition(s) in message base most closely matching proposition(s) in subject's protocol.

Table 3

Frequencies of Inference Types in Five Children's Recalls of Two Stories

Inference Type	Story 1					Story 2					Total
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	
Identifying Operations											
3A. Attribute Inference: States	1	1	1								
3B. Attribute Inference: Events											
4A. Category Inference: States						1					1
5A. Time Inference	5	4	6	2	2	1	3	1		3	27
Frame Operations											
10A. Frame Inference											
12A. Result Inference											
15. Theme Inference			1								1
16D. Frame Transformation						1				1	2
17C. Disembedding Operations: GOAL			1	1	1						3
17D. Disembedding Operations: THEME	1			1							2
18C. Embedding Operations: GOAL							1				2
18D. Embedding Operations: THEME			2								2

Table 3 (cont.)

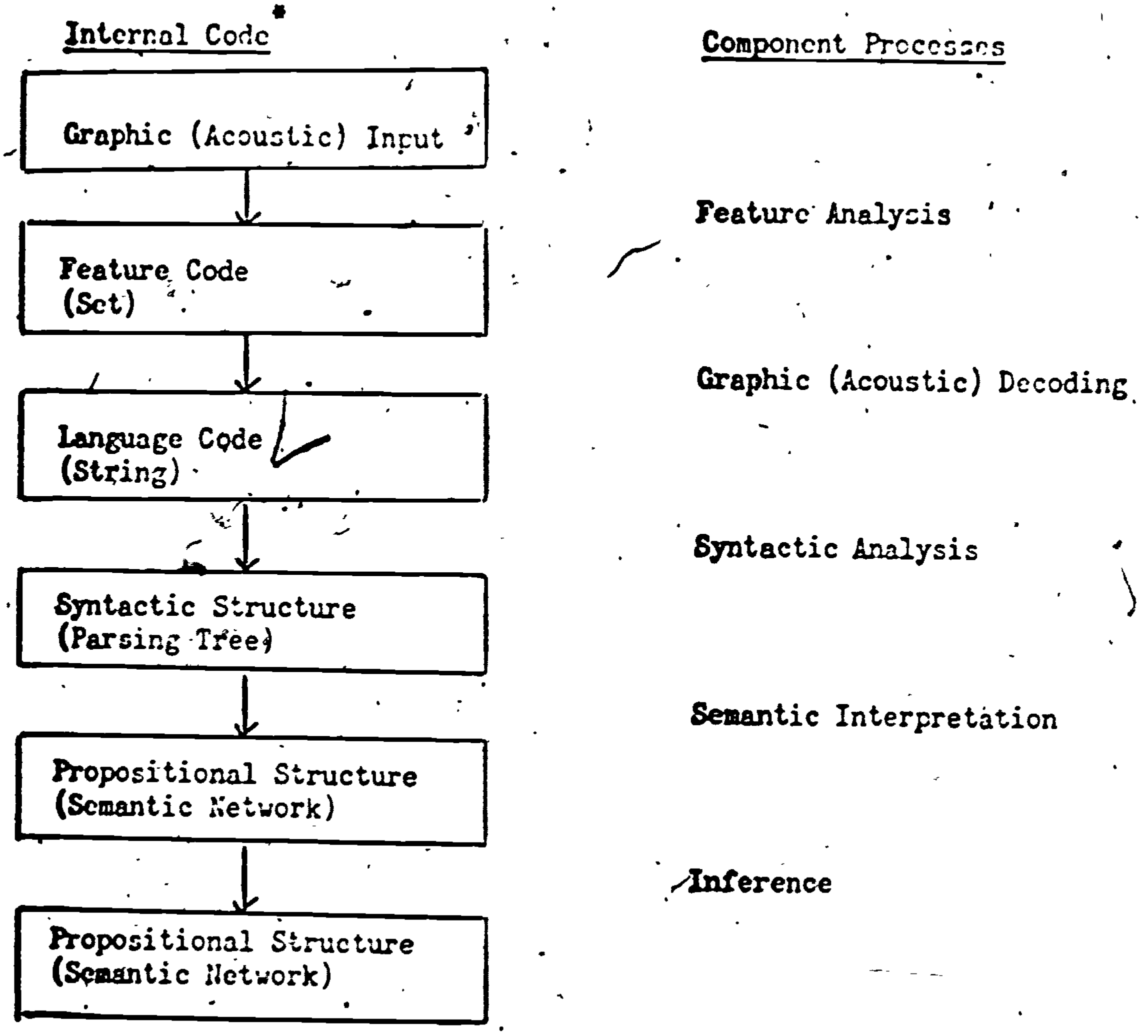
Inference Type	Story 1					Story 2					
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	Total
Event Generation											
Processive Event:											
19A PAT	1	1	1	1	1		1	2			8
19B OBJ2										1	1
19AB PAT + OBJ2			1								1
Resultive Event:											
19E AGT	1	1	2							3	8
19G OBJ1			1			3					4
19EG AGT + OBJ1							1				1
Macrostructure Operations											
Superordinate Inference:											
20B(AGT) Superordinate agent	1		1			2					4
20B(OBJ1). Superordinate object		1			1	1	1			1	6
20B(PAT) Superordinate patient						2				1	3
20E. Superordinate RESULT				1							

Table 3 (cont.)

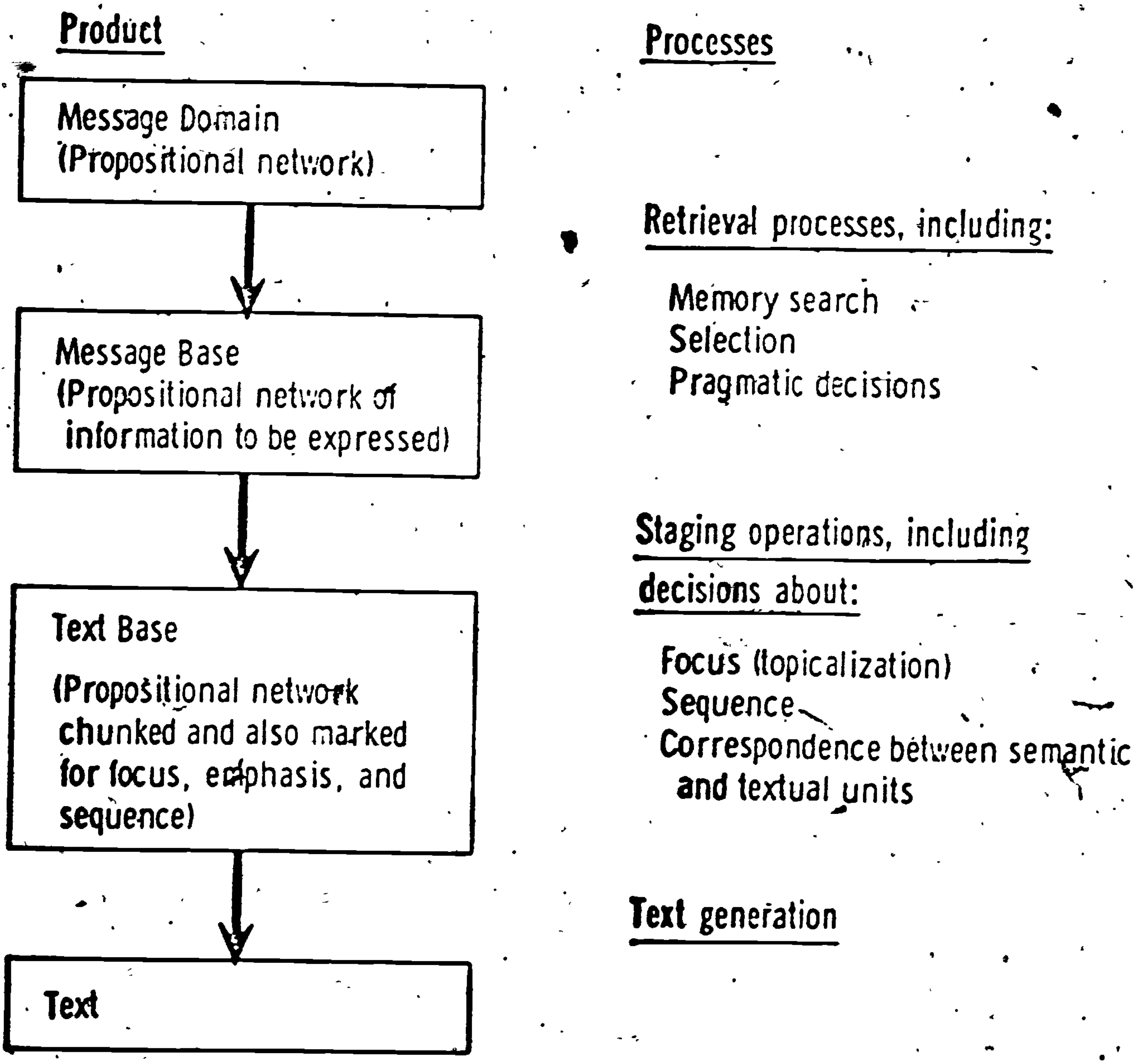
Inference Type	Story 1 S1	S2	S3	S4	S5	Story 2 S1	S2	S3	S4	S5	Total
Macrostructure Operations (cont.)											
Subordinate Inference:											
21B(ATT). Subordinate agent	1										
21B(OBJ). Subordinate object	1					1					2
21A. Subordinate action											
Algebraic Operations											
22A. Metric Algebraic Operations	2		2			2				3	9
22B. Nonmetric Algebraic Operations		1			1						2
Dependency Operations											
Causal Inference:											
23A. Connect unconnected events			1								
23F. Inchoative	1										2
Conditional Inference:											
24A. Enablement inference				1	1						2

Table 3 (cont.)

Inference Type	Story 1					Story 2					
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	Total
Truth-Value Operations											
26A. Qualification					1						1
26B. Negation	1										1
Subject Totals	17	9	20	7	8	15	7	4	[illegible]	18	



* Type of data structure indicated in brackets.


Product
Message Domain (Propositional network)
Message Base (Propositional network of information to be expressed)
Text Base
(Propositional network chunked and also marked for focus, emphasis, and sequence)
Text
Processes
Retrieval processes, including:
Memory search
Selection
Pragmatic decisions
Staging operations, including decisions about:
Focus (topicalization)
Sequence
Correspondence between semantic and textual units
Text generation

```
;---------------------------------------------------------------------------
;PROPOSITION LIST - MICHELE AND HER MOTHER
;---------------------------------------------------------------------------

100     (:MICHELE)--AGT;TEM(PRES);ASPCT(CONT)->('WATER)◆--OBJ1->
+(:FLOWERS)

200     ()--I.PAT->('CRASH)
201     (:MICHELE)--PAT;TEM(PRES)->('THINK)◆--THEME2->["202]
202     (:MOTHER)--PAT;TEM(FUT)->()--MAN.EXT1->('MAD)
203     (:MICHELE)--PAT;TEM(PRES)->('WANT)--GOAL2->["204]
204     (:MICHELE)--AGT->('RUN)◆--SOURCE->["205]
                                 /--RESULT->["206]
205     (:MICHELE)--LOC.0.2->(#205)
206     (:MICHELE)--LOC.0.2->(#206)
207     (#205)<-PROX;NEG->(#206)

300     (:MOTHER)--AGT;TEM(PRES)->('SAY)--THEME1->[["301],["302],
+["305]]
301     (:MOTHER)--PAT;NEG;TEM(PRES)->()--MAN.EXT1->('MAD)
302     (:MOTHER)--PAT;TEM(PRES)->('KNOW)--THEME2->["303]
303     (:MICHELE)--PAT;NEG;TEM(PRES)->('INTEND)--GOAL2->["304]
304     (:MICHELE)--AGT->()--RESULT->["200]
305     (:MOTHER)--AGT->('SUGGEST)--THEME1->["306]
306     ((:MOTHER),(:MICHELE))--AGT->('CLEAN)◆--OBJ1->(:MESS)

400     (:MICHELE)--AGT;TEM(PRES)->('PICK.UP)◆--OBJ1->(:FLOWERS)
401     (:MICHELE)--AGT;TEM(PRES)->('GIVE)◆--OBJ1->(:FLOWERS)
                                              /--DAT1->(:MOTHER)
403     (:MOTHER)--AGT;TEM(PRES)->('SAY)--THEME1->[["404],["406]]
404     (:MOTHER)--PAT->('WANT)--GOAL2->["405]
405     (:MICHELE)--PAT;NEG->('SORRY)
406     ((:MOTHER),(:MICHELE))--AGT;TEM(FUT)->('PUT)◆--OBJ1->(:FLOW
                                                      /--RESULT->["40
407     (:FLOWERS)--LOC.0.3->(:POT.TOP)
408     (:POT.TOP)--EXT0->('NICE)
```

```
500     ()←-AGT?TEM·PAST·?ASPCT(COMP)->('CLEAN)◆--OBJ1-->(:MESS)
501     (:FLOWERS)--CAT.ATT->('ON)
502     (:MOTHER)--AGT?TEM·PRES)->('SAY)◆--THEME1->
◆[["503]·["510]·["515]]
                                                   /--LOC.0·0->(#502)
503     (:MOTHER)--PAT->('WANT)--GOAL2->["504].
504     (:MICHELE)--AGT->('MOVE)--RESULT->["505]
505     (:MICHELE)--LOC.0·0->(#505)
506     (#502)<-PRO:->(#505)
510     ((:MOTHER)·(:MICHELE))--AGT->('MOVE)◆--SOURCE->["511]
                                                      /--RESULT->["513]
511     ((:MOTHER)·(:MICHELE))--L[illegible]0·0->(#511)
512     (#502)<-PRO:->(#511)
513     ((:MOTHER)·(:MICHELE)◆-LOC.0·0->(#513)
514     (#511)<-PRO:NEG->(#513)
515     ((:MOTHER)·(:MICHELE))--AGT->('MAKE)--RESULT->(:COOKIE)

$---------------------------------------------------------------------------
$OBJECT LIST - MICHELE AND HER MOTHER
$---------------------------------------------------------------------------

(:FLOWERS)      ('FLOWER)--DEF->()--NUM1->('PL)
(:MICHELE)      ('MICHELE)--DEF->()--NUM0->('1)
(:MOTHER)       ('MOTHER)--DEF->()--NUM0->('1)
(:MESS)         ('MESS)--DEF->()--NUM0->('1)
(:POT:TOK)      ('POT)--TOK->()--NUM0->('1)
(:COOKIE)       ('COOKIE)--TOK->()--NUM1->('SOME)
```

70

```
$-----------------------------------------------------------------------------
$PROPOSITION LIST - MICHELE AND HER MOTHER - SUBJECT 1 RECALL
$-----------------------------------------------------------------------------


100       (:GIRL)--AGT@TEM(PAST)@ASPCT(CONT)->('WATER)◆--OBJ1->
◆(:FLOWERS)


202       (:MOTHER)--PAT@TEM(FUT)->()--MAN.EXT1->('MAD)


306       ((:MOTHER),(:MICHELE))--AGT@TEM(FUT)->('CLEAN)◆--OBJ1->(:MESS)


403       (:MOTHER)--AGT@TEM(PAST)->('SAY)--THEME1->[*404]
404       (:MOTHER)--PAT->('WANT)--GOAL2->[*405]
405       (:MICHELE)--PAT@NEG->('WORRY)
406       (:MOTHER)--AGT@TEM(PAST)->('PUT)◆--OBJ1->(:FLOWERS)
                                           /--RESULT->[*407]
407       (:FLOWERS)--LOC.0,3->(:POT.TOP)
408       (:POT.TOP)--EXTO->('NICE)


500       ()--AGT@TEM(PAST)@ASPCT(COMP)->('CLEAN)◆--OBJ1->(:MESS)
501       (:FLOWERS)--CAT.ATT->('OK)


$-----------------------------------------------------------------------------
$OBJECT LIST - MICHELE AND HER MOTHER - SUBJECT 1 RECALL
$-----------------------------------------------------------------------------


(:FLOWERS)       ('FLOWER)--DEF->()--NUM1->('PL)
(:MOTHER)        ('MOTHER)--DEF->()--NUM0->('1)
(:MESS)          ('MESS)--DEF->()--NUM0->('1)
(:POT.TOP)       ('POT)--TOP->()--NUM0->('1)
```

OPEN DISCUSSION OF FREDERIKSEN PRESENTATION

SHUY: Carl, that is a very instructive and interesting paper I think, and I think we can benefit from it a great deal.

I am going to ask a question which is a lot like the one I asked about decoding a minute ago, because, from the perspective of my discipline, it seems that the word comprehension is being used in dozens of different ways in the field of reading.

I think that you define comprehension as ability to make inferences. Is that what you mean?

FREDERIKSEN: No. By comprehension, I mean whatever processing is involved in acquiring propositional information from discourse. Propositional information is of a variety of types, has a variety of relationships to a text, and involves a variety of processes interacting at different levels.

SHUY: So, inference, as you use it is really a key for a number of other kinds of meanings that could be used?

FREDERIKSEN: Right. I restricted my use of the word inference to refer to processes that operate solely at a propositional level. Incidentally, those who use the term decoding also don't mean a simple unitary process. I think Dick Venezky indicated that it's a very complex process. These are just convenient labels that we use to refer generally to the domain of processing that we are talking about.

SHUY: I don't deny that it's possible to have a technical meaning of the term, used by a given field, but when we get into any kind of cross-field suggestions, terms such as comprehension get a little fuzzy. We would use it in linguistics to describe semantic meaning or pragmatic meaning.

FREDERIKSEN: I think that both are related to what we are discussing. Inference is very closely related to pragmatic meaning and presupposition. For example, I have been working on a study of children's conversations with Drs. John Dore and William Hall at Rockefeller University in which we are attempting to code both illocutionary functions of speech acts and relations among propositions. What we are trying to do is look at conversations both from a propositional point of view and from the point of view of the illucationary functions of speech acts. The notion is that it should ultimately be possible to develop a description of the cognitive processes which operate in conversations and underlie the semantic, functional, and pragmatic relationships among speech acts.

SHUY: When you talk about the cognitive processes involved in comprehension, can I assume that you mean, and other people in reading mean, those hidden meanings that are so common in the Stanford reading achievement tests? Do you mean, for instance, some things that really look like problem solving, where you are given a task to do, and you have to go back and search through the text to see whether there were seven yellow rabbits running down the road, and, then, perhaps find some other definitions as well? Those tests always confused me, becsause it looks like they lump all of these kinds of things together and give you a score.

FREDERIKSEN: Yes, comprehension includes certain kinds of problem solving. Such global scores are meaningless to me; it's like counting up all of the animals in the barnyard. You don't know what you have when you're through counting.

SHUY: Aren't you really, however, referring to all of those kinds of tasks?

FREDERIKSEN: Yes. While we haven't yet a detailed account of the comprehension process, I do believe we have a way of identifying the different aspects of comprehension. In fact, while the study of comprehension is becoming theoretically very sophisticated, the data on which theories are based is very thin. I have been trying to improve this situation by developing empirical techniques which can provide a reasonably rich source of information about inferential processes in comprehension.

GREGG: You said that all of this is problem solving. What about an algebra word problem? You know, it's English.

FREDERIKSEN: It involves comprehension.

GREGG: When does it become comprehension, and when does it become problem solving?

FREDERIKSEN: I think comprehension is a kind of problem solving.

GREGG: All the way?

FREDERIKSEN: It just depends on what you want to call it.

I don't have a good definition of problem solving, nor do I see these distinctions as clear-cut. All I want to do is describe what the human cognitive system does in acquiring propositional knowledge from text. This can involve operations on propositions which may be described as problem solving.

GREGG: Well, I think that's part of the problem that we are all having at the conference. We are just going to call everything by one big global name, so we all have a chance to think in our own way about it.

FREDERIKSEN: Yes, I agree that is a problem. I have tried to facilitate the communication by distinguishing the several different kinds of internal codes. For example, one can talk separately about a decoding component by referring to processes which encode information in terms of an abstract language code. The problem here is that the inference category is exceedingly large. So far, I have been dealing with the lowest levels of inference, but, there are clearly very complex heuristics underlying such inferences in discourse comprehension, heuristics that we eventually want to describe. My strategy is to "ease into" the problem from the bottom up. However, I have been attempting to describe text-based inferences in great detail, not in terms of a few global categories.

GREGG: Well, I got Roger to admit there was a decode one and decode two and decode three, so there must be several technical terms that we could come to some agreement on.

SAMUELS: What decodes are you referring to?

GREGG: Oh, going from letters to words, from orthographics to the words themselves, but not from words to meanings. Those are the things we really mean by decoding. In your case, we mean going from the sensory code to the language code, which is going from one code to another code, and so maybe it isn't decoded yet.

FREDERIKSEN: I will be happy with any terms that we can mutually agree on.

KINTSCH: I think it would help, for instance, in this first slide of yours, if you would add at least one other level at the bottom of that slide. You left out meaning at the propositional level.

Now, what Ken was talking about right before you brought up the different levels was a functional level, which I think would fit in very well if you would add the meaning level.

FREDERIKSEN: The trouble is that I don't know what internal code other than a propositional code is associated with function.

KINTSCH: That doesn't really make any difference; you don't really know what is associated with the other things either. You know it exists, though.

BLOCK: I think that your work is interesting, Carl, in terms of its potential for providing a way to refine and elaborate our taxonomy of comprehension skills, and I think it would be very useful for some of us who have an interest in that to take a look at the comprehension instruction in beginning reading programs. That is something Isabel and I did not have time to do. I am certain that there is some overlap between the knowledge and skills that programs develop and those that are required, but we don't know the degree to which there is exact overlap.

We really do not know the extent to which the things that need to be learned are actually being taught, although these things may not be explicitly stated as being taught.

FREDERIKSEN: Yes, I would be very interested in that; I think that would be a very worthwhile project. It would also be interesting to see how it relates to the often expressed remark that there is a difference in the kind of reading that children are required to do in the early grades and what happens in the middle grades. I heard that said a lot, and I have never seen any kind of detailed analysis of the change that takes place, both in the textual materials and in the context, including functional context in which they are used.

SINGER: I want some clarification on your bottom-up and top-down processing.

Don't you always go from bottom to top and then from top to bottom? You don't start at the top, in your top-down, do you?

FREDERIKSEN: Well, I was starting with the assumption that you have a language code, first of all. But one always starts at the bottom. The question is how far do you go, so that Shank just goes a little way, and then he starts doing things at a higher level. That's the reason I said it is an oversimplification to talk about a bottom-up model versus a top-down model. I think this is just a way of characterizing differences that can occur in the way component processes interact.

You should see, for example, the flow chart MIT has put out, a thing called "People's Flow Chart for the Winograd Program." It's incredible. Have you ever seen it? It's just on the interaction, the way those subroutines interact.

There is no simple distinction between a parser and a semantic interpreter and a logical component. The Winograd approach to grammar is called procedural grammar--a series of programs or procedures which ultimately generate semantic representations not unlike the one that I am using.

DANKS: Could you expand on your comment that you felt that the oral reading task biased the child to a bottom-up approach?

FREDERIKSEN: I don't know this to be true, but it seems a reasonable guess. You could find out by studying the kinds of information children acquire from a text read aloud or silently. You could have a child read aloud and then have the child retell the story and compare the child's recall to that obtained by just asking the child to read a story, or listen to a story. One would look at the kinds of information the child incorporates into his retelling of the story under the different conditions.

One effect of oral reading is to increase reading time. Oral reading would also constrain the way in which a line of text is scanned. For example, suppose eye fixation data were studied in oral reading as opposed to silent reading. If one found major differences, I would assume that would be evidence for different processing strategies. You would probably find more fixation points on a line on a text in oral reading than in silent reading. Is that right?

CHALL: That's on a level.

DANKS: Because the reader is going slower.

FREDERIKSEN: A level of reading, you mean?

CHALL: A level of reading.

DANKS: But the oral reading rate slows the reader down, there would be a constraint on fixations.

FREDERIKSEN: You could maintain the same number of fixation points and increase fixation durations, for example. As a matter of fact, the most interesting work on eye movements now, I think, looks at eye movements from the point of view of the cognitive processes that control them.

DANKS: I asked the question because there is some suggestion in the literature that, in the oral reading task, the reader does comprehend, and the oral output is generated from a higher semantic representation. I don't know exactly what the nature of that semantic representation is.

FREDERIKSEN: It seems likely that oral reading can occur with or without comprehension. I can read stories that I have read many times before to my son without processing the stories at the semantic level at all. In fact, I can even be thinking about a problem while I am reading aloud to my son, or I can be thinking about the story. My son can recognize whether I am thinking about the story or not. He'll say, "You are not paying attention, Daddy."

PRESENTATION BY DENNIS FISHER

RESNICK: We move next to Dennis Fisher's paper, "Dysfunctions in Reading Disability: There's More than Meets the Eye."