Discourse comprehension involves readers or listeners in constructing mental models using local text information, global text structures, and their general knowledge of the world. An analysis of the literature on children's understanding of spoken discourse reveals that young children are capable of forming mental models from texts, but that their overall level of comprehension may be reduced by factors such as limited general knowledge, inexperience in constructing certain types of mental models, unfamiliarity with particular global text structures, difficulty in understanding anaphoric expressions, and constraints on memory. Differences between written texts and the spoken discourse that children have mastered when they begin school have the potential to exacerbate their comprehension problems. Written text taps a wider range of general knowledge, shows different forms of discourse organization, uses different anaphoric devices, and provides less contextual support than spoken discourse. The major hurdle in the acquisition of discourse comprehension skills may be expository prose. Since more of the knowledge that students are expected to acquire in school is conveyed in that format, facility in dealing with expository discourse represents a crucial step in the development of text understanding. (HOD)
Technical Report No. 343
DISCOURSE STRUCTURE AND MENTAL MODELS

J. Kathryn Bock
Michigan State University

William F. Brewer
University of Illinois at Urbana-Champaign

November 1985

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

The work upon which this publication is based was performed pursuant to Contract No. 400-81-0030 of the National Institute of Education. It does not, however, necessarily reflect the views of this agency. The authors thank Harry Blanchard, Anne Hay, and Stella Vosniadou for their comments on an earlier draft and Thomas Carr for meticulous editing.

Discourse Structure and Mental Models

Abstract

In this paper we argue that text comprehension can be viewed as the process of constructing mental models from texts. We suggest that the construction of mental models involves the use of local text information, global text structures, and the reader's general knowledge of the world. We analyze the literature on children's understanding of spoken discourse and conclude that young children are capable of forming mental models from texts, but that their overall level of comprehension may be limited by factors such as limited general knowledge, inexperience in constructing certain types of mental models, unfamiliarity with particular global text structures, difficulty in understanding anaphoric expressions, and constraints on memory. We contrast the comprehension of spoken discourse with the comprehension of written text. We conclude that written text taps a wider range of general knowledge, shows different forms of discourse organization, uses different anaphoric devices, and provides less contextual support than spoken discourse. Finally we discuss the implications of this analysis for reading in terms of the transfer of oral comprehension skills to the understanding of written text. We suggest that the transfer should be relatively easy for narrative texts but more difficult for expository texts.
Discourse Structure and "Mental Models"

This paper focuses on the discourse features of text and their influence on comprehension and learning to read. Our point of view is that discourse comprehension involves the construction by readers or listeners of mental models synthesized from the information in the text and their general knowledge. Within this framework, we examine five aspects of discourse comprehension, first in the context of young children's ability to understand spoken discourse, and then from the perspective of differences between speech and writing. These five aspects include the use of knowledge, the kinds of mental models that underlie different types of discourse, the global organization of texts, the use of discourse cues in the construction of mental models, and the integration of information into a mental model. We conclude by exploring some potential implications of this approach for the development of reading skill.

Mental Models and Discourse Comprehension

Mental models. Watching a movie based on a familiar novel often brings on the feeling that something is awry. The rooms are too large, the furniture too new, the protagonist too handsome. Such a feeling presumably grows out of the contrast between what actually unfolds on the screen and expectations built on earlier imaginings about the people, places, and events of the story. We assume that these imaginings reflect a
Discourse Structure and Mental Models

fundamental part of discourse comprehension, which involves the construction of a mental model. Mental models are mental representations of particular states of affairs, such as events or places or someone's wishes. In discourse comprehension, listeners and readers try to construct mental models that embody the content of the text (Collins, Brown, & Larkin, 1980; Johnson-Laird, 1980, 1983).

For the listener or reader to construct a mental model, the sentences in a discourse should be coherent and describe a plausible set of ideas or sequence of events (Johnson-Laird, 1983). Coherence depends in part on coreference among the sentences of the text: For the comprehender to construct a single integrated mental model, every sentence must directly or indirectly refer to something mentioned in another sentence. Plausibility requires the discourse to be interpretable within a unified framework consistent with the comprehender's knowledge of time, space, causation, and human intention. It is possible to compose a passage that is coherent but not interpretable within a unified framework; for example:

 Robbie owned a bike. It was made in Great Britain. Great Britain is an island. On the island are several monolithic structures. These structures may have been early astronomical observatories. Modern observatories use telescopes.
However, coherence and plausibility are ordinarily associated in natural discourse.

This analysis yields a reasonable straightforward approach to the process of discourse comprehension. To understand a coherent, plausible discourse is to construct a mental model of the events, descriptions, or arguments that underlie it, integrating one's general knowledge and the information in the text into a unified representation.

**Knowledge.** Even the very young comprehender brings an enormous amount of real-world knowledge to the task of language understanding, knowledge that we assume is represented in the form of generic structures called schemas (Brewer & Nakamura, 1984; Minsky, 1975; Rumelhart, 1980). To the degree that a segment of discourse makes contact with the comprehender's schema-based knowledge, the information can be used to construct a mental model that is much richer than the information explicit in the text. Even linguistically impoverished prose, such as *First bike ride--man--boy--push--wobble--peddle--fall--grin* can be readily interpreted by most adults, using knowledge of first bike rides to construct a tentative model that goes far beyond the text. Schema-based knowledge thus powerfully influences the representation that is developed as spoken or written discourse is understood.

A classic demonstration of the influence of schema-based knowledge on comprehension supports the claim that adults'
interpretation of text yields an integrated representation that includes more than the presented information. Bransford, Barclay, and Franks, (1972) showed that subjects who heard such a sentence as Two turtles rested on a floating log, and a fish swam beneath it often thought they had heard the test sentence Two turtles rested on a floating log, and a fish swam beneath them. They apparently inferred that the fish swam beneath the turtles as well as the log, although this was not directly stated. When the original sentence was Two turtles rested beside a floating log, and a fish swam beneath it, subjects much less often claimed to have heard the test sentence (see also Garnham, 1981). These patterns of false recognition can be explained by the assumption that listeners constructed mental models of the state of affairs that the text was intended to convey.

**Discourse information.** It is obvious that mental models are not completely inferential. The other major source of information used in constructing them is the explicit language of the discourse. This language details the setting, identifies characters, describes specific events, or lays out the ideas and arguments that the author wishes to make explicit. Some of the interactions between a developing model and the language of a text are illustrated in the following narrative about a first bike ride:
One cloudy October day, Fred Bartlett took his son Robert out for his first bike ride. Robbie was very excited and said, "I sure am glad I got this Raleigh for my birthday." His father smiled and checked to see if the football field was clear. He picked Robbie up and put him on the new two-wheeler. His son was trying to be brave and had a very serious look on his face. Mr. Bartlett gave him a push, and the secretly terrified child began to peddle. He wobbled briefly and then went straight about fifty feet before he fell over. His father ran over and found a little boy lying on the grass with a big grin on his face.

Consider the differences between the mental models that a typical adult might produce for the previous minimal text and this extended narrative. Because little specific information is given in the minimal text, most of the model must be generated from schema-based knowledge, and seems likely to include the information that the bike is a bicycle (not a motorcycle), that the older person instructs the younger, that the older person is the father of the younger, and that the younger person falls off the bike. In these instances, the schema-driven inferences are confirmed by the language of the full text. However, a mental model for the minimal text might also place the ride on a sidewalk or street on a warm sunny day. These inferences are not supported by the extended narrative. The information that the ride occurred on a football field on a cloudy October day, which
should be represented in the mental model generated from the passage, comes from the text.

Certain linguistic features serve as cues for specific types of interactions between the information in the text and the developing mental model of the comprehender. In the preceding narrative, two individuals and a hike are mentioned frequently, but the expressions denoting them vary. The boy is referred to as his son Robert, Robbie, I, him, his son, the secretly terrified child, he, and a little boy, and his father is referred to as Fred Bartlett, his father, he, and Mr. Bartlett. The reader must correctly interpret each of these expressions, determining which person is intended. Part of the linguistic information can be used to figure out whether an expression denotes something that is already part of the mental model (given information) or something that is to be added (new information).

Reference to given information is often marked by pronouns (I, him, he) or other anaphoric devices, such as definite noun phrases like the secretly terrified child. To understand the text adequately, the appropriate information must be located in the mental model and modified by adding the new material to the representation (Clark & Haviland, 1977). New information more often conveyed in full noun phrases than in pronouns, and those noun phrases may be marked with an indefinite determiner (for example, a push, a big grin) (Halliday & Hasan, 1976). In the last sentence of the passage, a is used inappropriately to refer
to information already in the reader's mental model (a little boy); such inappropriate marking seems to disrupt reading (Irwin, Bock, & Stanovich, 1982) and the integration of information into a unified representation (deVilliers, 1974).

These are some of the basic features of mental models and their construction. We turn now to the issue of children's ability to create mental models to represent the information in spoken discourse.

**Children's Understanding of Spoken Discourse**

A number of experiments suggest that young children can construct mental models from spoken discourse. Their ability to integrate linguistic information with relevant knowledge is especially clear in a study by Brown, Smiley, Day, Townsend, and Lawton (1977). In the second experiment of that study, second graders heard a narrative about a hunter of the fictitious Targa tribe. Although the story contained no information about weather, climate, or terrain, the children had heard a passage about the Targa a week before that described them as Eskimos living in a cold climate or as Indians living in a desert. In response to questions about weather and terrain, most of the subjects not only answered in accordance with the information acquired the week before but also said they were sure it was part of the story they had just heard. This suggests that they integrated the narrated content into a mental model incorporating prior knowledge. In other research children between the ages of
Discourse Structure and Mental Models

four and six have been found to understand the intentions of characters and causal relations among events (Stein & Glenn, 1979) and to notice incongruities in stories violating schema-based expectations about events, such as the theft of a bicycle or losing money (Stein & Trabasso, in press; Wimmer, 1975), even though these implications were not spelled out in the stories that they heard.

Despite young children's ability to construct mental models from spoken discourse, there are apparent limitations on their performance. The second graders in the study by Brown et al. (1977) seemed to develop less elaborate representations than those of older children who performed the same tasks. Omanson, Warren, and Trabasso (1978) found that five-year-old children made many fewer inferences about the implicit content of stories than eight-year-olds did. Elementary school children in experiments by Markman (1977, 1979) regularly failed to detect omissions and inconsistencies in instructions and prose passages, leading Markman (1981) to suggest that they tended to treat the individual statements of texts as isolated units instead of constructing integrated representations. Other investigators have shown that children between the ages of five and seven often do integrate the information in a passage well enough to recognize implied relationships accurately (Liben & Posnansky, 1977; Hoesser, 1976; Paris & Upton, 1976; Small & Butterworth, 1981), even when the separate items needed to make the correct
Discourse Structure and Mental Models

Inferences are available in memory (Collins, Wellman, Keniston, & Westby, 1978; Schmidt, Paris, & Stober, 1979). As a result, a young child who has heard a short descriptive passage, such as

The bird is in the cage. The cage is under the table. The bird is yellow.

may not integrate the sentences into a mental model carrying the information that the bird is also under the table (Small & Butterworth, 1981).

Because these problems are most striking in preschool and young school-age children, they may be relevant to children's ability to understand text in the early stages of reading. However, the explanation for these problems is far from clear: There is no immediately obvious reason why children understand spoken discourse so well on some occasions and so poorly on others. We will discuss five possible sources of difficulty, each representing a different aspect of the construction of a mental model from spoken discourse, before we consider the further complications that written texts present.

Knowledge. Young children lack some of the general knowledge that older children and adults possess (Chi, 1978). The importance of such knowledge in language comprehension and memory has been demonstrated repeatedly. For example, Bransford and Johnson (1973) report a study in which sentences like The notes were sour because the seam split were found to be hard to recall—unless they were preceded by a word, such as bagpipes,
that cued pertinent knowledge. In the experiment by Brown et al. (1977), children who had acquired relevant knowledge a week earlier remembered more of the story than children who had been given irrelevant information. Because of their inexperience and consequent lack of knowledge in many domains, younger children may construct relatively impoverished mental models, or fail to construct them at all.

Types of mental models. Brewer (1980) has argued that different types of mental representations underlie texts from various genres. He proposed that descriptive discourse is represented by visual-spatial structures, narrative by plan and event structures, and expository text by abstract propositions or thoughts. Some of these representations may be easier to construct than others. Along these lines, several researchers have suggested that narratives are easier to understand than expository texts (Bereiter & Scardamalia, 1982; Spiro & Taylor, in press). In the experiment by Markman (1979) cited earlier, elementary school children failed to notice inconsistencies in expository passages, although in other experiments much younger subjects were able to detect incongruities in narratives (Stein & Trabasso, in press; Wimmer, 1979). One possible explanation for this disparity is that it is harder to form a mental model for the abstract logical structures that underlie expository texts than it is for the actions and events portrayed in narratives.
Global discourse organization. The information in a particular type of discourse is often organized in a characteristic way. For example, in folktales from the oral tradition, information about the characters, time, and location of the story ("Once upon a time in a land far away, there was a princess . . .") typically precedes the recounting of the events (Propp, 1968). Nonfictional spoken narratives commonly begin with setting information ("Last Thursday, Robbie was riding his bike on the street . . .") (Chafe, 1980). Oral narrations of events regularly follow the chronological order in which the events occurred, and the elements of spoken descriptions tend to conform to the order in which things are encountered in a spatial layout (Clark & Clark, 1968; Levelt, 1981; Linde & Labov, 1975; Osgood, 1971).

Although many of these conventions appear to be very natural, alternative forms of organization are possible, and they are used. Certain of these alternatives have been found to disrupt younger children's language comprehension and memory. For example, Stein and Nezworski (cited by Baker & Stein, 1981) changed the order of mention of events in a narrative relative to the order of occurrence, marking the inversions in a way that indicated the deviation (for example, "Robbie broke his leg. It happened because he rode his bicycle into a parked car" instead of "Robbie rode his bicycle into a parked car. He broke his leg"). Although fifth graders recalled the information at least
as well when it was conveyed in marked inversions as when it was normally ordered, first graders recalled some types of information less well from narratives in the inverted format. It seems that discourse forms whose organization departs from the conventional structure or discourse forms whose conventions are unknown to children may impede comprehension, either because of their cognitive complexity or because children lack appropriate discourse knowledge.

**Using discourse cues.** Children's difficulties with text integration may in some cases be traced to inefficient processing of anaphoric devices that mark repeated reference, such as pronouns. In an experiment by Tyler (1983), adults and five-, seven-, and ten-year-old children listened for mispronunciations of words in a spoken text. The mispronunciations were strategically located after pronouns or definite noun phrases that were coreferential with an expression in the preceding sentence, as in the following examples:

Mother saw the postman coming from a distance. **He** brought a letter from Uncle Charles who lives in Canada.

Mother saw the postman coming from a distance. **The postman** brought a letter from Uncle Charles, who lives in Canada.

The referent of the italicized expressions should have been part of the listener's current model, making letter (mispronounced leffer) contextually predictable. However, five-year-olds were slower to detect the mispronounced word following the pronoun
than following the definite noun phrase, although adults and
older children showed no differences between the two conditions.
This suggests that the younger children had more difficulty
accessing the referent of the pronoun from their representation
of the content of the preceding sentence, which slowed their
integration of the sentence into the developing mental model.

Information integration. The final explanation that we will
consider for disruptions in the formation of mental models is a
general information-processing problem. Inferring relationships
among the elements of a text demands that the relevant pieces of
information (from the mental model, general knowledge, or
immediate discourse) not only be stored in memory but also be
actively in mind—held in working memory—at the time when the
inference is to be made (Hayes-Roth & Thorndyke, 1979; Walker &
Meyer, 1980). Children may be able to keep less information
active in memory because of inefficient use of working memory
capacity (Case, Kurland, & Goldberg, 1982), knowledge limitations
(Chi, 1976), or diminished memory capacity (Pascual-Leone, 1970).
They may also neglect to retrieve and represent information in
working memory when it is needed for integrating new material. A
large body of evidence shows that children before roughly the age
of seven do not spontaneously employ memory storage and retrieval
strategies commonly used by older children and adults. Although
younger children are able to use such strategies when instructed
to do so, and although they benefit when they use them, they do
not invoke them without prompting (Brown, Bransford, Ferrara, & Campione, 1983).

When measures are taken to alleviate memory-related problems, children's ability to draw inferences also improves dramatically: Four-year-olds have been found to infer unseen relationships among objects (Bryant & Trabasso, 1971) and unstated relationships among things mentioned in sentences (Harris & Basset, 1975) nearly as well as adults. However, without some form of external support for the active maintenance of appropriate information in memory, young children's inferences about relationships among the elements of spoken discourse may be restricted.

**Summary.** We have suggested five aspects of the construction of mental models from discourse that represent possible problem areas in comprehension for preschool children—children who are about to begin learning to read. Thus far, however, we have focused on research concerning the understanding of spoken discourse. Because spoken and written language differ in substantial ways, the child confronted with a written text needs new solutions to some of the foregoing problems. In the next section, we consider the changes that occur in the transition from listening to reading.

**Spoken Versus Written Discourse**

Spoken and written discourse differ on a number of dimensions (Brewer, 1985; Chafe, 1982; Rubin, 1980; Snow, 1983;
Tannen, 1982). The dimensions that are most relevant for our purposes are interaction, contextualization, and transience.

Interaction refers to the mutual determination of form, content, direction, and pace of communication by the participants in an exchange. Encounters between people using spoken language typically include a speaker and an addressee who can respond to one another, make comments, and ask questions, and who do so under the constraint of contributing fairly rapidly or risking the loss of a turn at speaking. In written language, the writer receives no immediate feedback, but is solely responsible for shaping the discourse and unbothered by interruptions from the intended audience. Contextualization involves the sharing of spatial and temporal contexts. Speakers and listeners are often in the same place at the same time, but writers and readers are not: Written messages are usually produced and understood in different contexts. Finally, speech signals are transient: They are generally available to the listener only briefly. Written language is relatively permanent. As a result, it can be read a number of times, and the reader can refer back to previous text when necessary.

These and other, related dimensions create differences between spoken and written discourse that have implications for reading comprehension. We will discuss some of these differences in the context of the five facets of discourse comprehension introduced in the preceding section.
Knowledge. An oft noted advantage of the development of a writing system is that it allows the accumulated knowledge of a people to be passed from generation to generation in a form that is less subject to distortion and loss than oral transmission is. One obvious consequence of the accumulation of knowledge in print is that a wider range of topics and greater depth of coverage may be found in the books of an elementary school library than in day-to-day encounters with spoken language. Written discourse thus draws on and adds to a more diverse knowledge base than spoken discourse typically does.

Some evidence that the possession of specialized knowledge can contribute to discourse comprehension comes from an experiment by Spilich, Vesonder, Chiesi, & Voss (1979). They compared the ability of subjects who varied in their knowledge of baseball to recall and answer questions about an account of a portion of a baseball game. High-knowledge subjects recalled more and answered more questions correctly, and proportionately more of the information that they remembered pertained to major points from the passage. In general, people with more knowledge about a subject may be better at relating new information to old because information from a familiar domain can be maintained in active memory more efficiently than unfamiliar information can (Chase & Simon, 1973).

Types of mental models. Some types of discourse, including narrative, tend to occur in both spoken and written discourse.
Other types of discourse, such as exposition, are more commonly written than spoken. Certain written texts may thus require the creation of mental models that are both different and more difficult to construct than the models used to represent most spoken discourse.

Whether it is the difficulty of constructing its underlying representation or some other factor, expository prose appears to slow down even skilled readers more than narrative does. Graesser, Hoffman & Clark (1980) found that college students read narrative passages faster than expository passages—on the order of 140 milliseconds per word faster—and that narrativity was by far the best predictor of variations in reading time in analyses that also examined effects of topic familiarity, number of words, syntactic complexity, number of propositions, and number of new referents introduced in the text. The significance of the problems created by expository prose can be appreciated by considering the amount of information that students are expected to learn by reading expository texts.

**Global discourse organization.** The information conveyed in a discourse of a particular type may be organized in various ways. For example, rather than opening with a setting as oral narratives do, modern written stories tend to open with an event (O'Faolain, 1951). Setting information is instead woven into the text. Such variations in global discourse organization may play a larger role in writing than they do in speaking. The press of
time, limitations of memory, and the interruptions of interlocutors often constrain speakers' ability to organize a message. However, the author of a written text can organize and structure discourse over a longer period, even planning a 300-page narrative in which the reader receives information that forces a complete revision of the mental model on the last page. With expository text, a writer can organize a complex set of logical relations in text form and use headings and other structural marking devices to delineate them (Bereiter & Scardamalia, 1982; Chafe, 1982).

These and other differences between the usual circumstances and products of talking and writing may have increased the number of conventionalized global organizations in printed discourse. Brewer (1985) has claimed that written genres have a greater number of specialized text structures (newspaper articles, psychology journal articles, comic books, cookbooks, and so forth), each with its own conventions of content and form. Because different text organizations deal in different ways with the problem of presenting underlying cognitive structures in a sequential linguistic format, readers may benefit from a complementary inventory of comprehension strategies that are more varied than those used in listening. For example, because the pyramid style of newspaper writing summarizes important points before addressing the material in detail, readers who understand this organization can easily skim the material if they choose.
Using discourse cues. The absence of immediate conversational feedback and shared time and place requires written discourse to be more explicit than spoken. Chafe (1982) has provided evidence that writers pack more information into segments of text than speakers do. One of the important functions that this additional information serves is ensuring that readers correctly identify intended referents in their mental models. When speakers and hearers share the same context, simple expressions suffice to indicate the topic of an utterance. In written language, more cues are needed: A speaker in a conversation might nod his head and say over there to convey the same information as the decontextualized The old man they had seen earlier walking his Saint Bernard came into view across the street.

A related consequence of the contextualization of spoken language is that certain uses of pronouns and other referring expressions are more common in speech than they are in writing. The use of referring expressions to point out elements of the extralinguistic context is called deixis, while their use to indicate elements of a mental model that has been formed from a text is called anaphora. Someone watching a boy who has just hurled several objects at a wall might say to a companion, with no prelude, "What do you suppose he was doing?", where he is used deictically to indicate the boy. The same sentence in a written text with no introduction is cryptic. Instead, a referent is
usually established before the pronoun is used, as in "I saw a boy hurl several objects at a wall. What do you suppose he was doing?". Here, he is used anaphorically to indicate something that the reader or addressee should have in mind as a result of understanding the prior discourse. Although deixis is possible in writing, it is much more frequent in speech.

Determining the referent of a deictic pronoun typically requires the identification of a salient element in the extralinguistic context or the current focus of attention. Understanding an anaphoric pronoun demands careful examination of the characteristics of the pronoun and its syntactic role in addition to an evaluation of the characteristics of candidate referents in the mental model. If readers attempt to understand anaphoric uses of pronouns in the same way that listeners understand deictic uses of pronouns, perhaps by picking out the most salient elements of their current mental model, they may be unsuccessful in determining the correct referents. There is some evidence that less-skilled readers approach anaphoric pronouns in this way. Frederiksen (1981) compared less-skilled and better high school readers' ability to recover the antecedents of pronouns, and found that the less-skilled readers relied more heavily on a salience strategy. This strategy involved falling back on the topic of the passage as the referent. As a result, less-skilled readers read sentences containing pronouns with nontopical antecedents more slowly, and they were less likely to
identify the antecedent correctly than they were when the pronouns had topical antecedents. Such a pattern suggests that Frederiksen's less-skilled subjects may have dealt with pronouns in reading in a manner more appropriate to listening.

There is an additional distinction among the ways in which pronouns are used that has potential implications for understanding discourse cues in written language. This is the contrast between deep and surface anaphora (Hankamer & Sag, 1976; Webber, 1980). Certain types or instances of anaphora require a representation of the actual language of an earlier sentence to be recovered in order to understand the reference, while others, like those we have been considering, refer directly to nonlinguistic elements in the comprehender's mental model of the text. Anaphora of the former type, called surface anaphora, is less acceptable when sentences intervene between the anaphor and its antecedent. Compare these two examples of surface anaphora from Tanenhaus, Carlson, and Seidenberg (1985):

Somebody has to paint the garage. The paint is peeling and the wood is beginning to rot. Let's take a vote and see who.

Somebody has to paint the garage. Let's take a vote and see who.

It is more difficult to interpret who as who has to paint the garage in the first example than it is in the second, where there is no intervening sentence. A plausible explanation is that the
explicit linguistic representation of the initial sentence is no longer recoverable: After reading or hearing a sentence in discourse, readers and listeners have been shown to experience considerable difficulty in remembering the surface structure of prior sentences (Change, 1980; Jarvella, 1971). However, with deep anaphora, interruptions are less disruptive, since the reference is to a component of the mental model:

Somebody has to paint the garage. The paint is peeling and the wood is beginning to rot. Let's take a vote and see who has to do it.

If a reader is unable to remember the surface structure of a prior sentence, he or she can usually read it again. Listeners do not have this option. The transience of spoken language may thus lead speakers to use surface anaphora less often than writers do, creating another type of discourse reference for readers to master.

Auditory and visual presentations of language have other subtle effects on discourse cues. Spoken English depends on intonation as a primary indicator of givenness and newness, with new information typically receiving higher stress than given information. Beyond such conventions as underlining for emphasis, written language possesses few means for indicating variations in intonation. Instead, skilled writers rely on syntax to mark distinctions between given and new information, placing new information later in sentences than the given
Discourse Structure and Mental Models

Information to which it relates (Smith, 1971). Bolinger (1957) has claimed that this organization capitalizes on readers' generation of implicit intonation contours in which the highest stress is located near the ends of clauses. In listening, intonation seems to influence adults' cross-sentence integration, while syntactic variations that may be used to distinguish given and new information have little effect (Bock & Mazzella, 1983). In contrast, structural variations do influence integration in reading (Yekovich, Walker, & Blackman, 1978; also compare experiments 1 and 2 with experiment 3 in Ehrlich & Johnson-Laird, 1982). These findings suggest that readers use synt. more heavily than listeners do for discourse cues, either for direct indicators of givenness or newness or for indirect cues mediated by implicit intonation.

Information integration. Because it is easier to integrate material from separate sentences into a coherent mental model when the items of information to be related are simultaneously active in memory, conditions that increase the probability of concurrent activation should enhance integration. For example, Walker and Meyer (1980) found that adult readers integrated text information more often when the separate components occurred consecutively than when they were separated in the text.

However, only a subset of the information from a text will be readily accessible at any one time, because the amount of information from a discourse that can be activated simultaneously
is limited. Kintsch and van Dijk (1978) hypothesize that these limitations are reflected in the number of propositions from a text that can be maintained in working memory. We assume that the comprehender must interpret these propositions in terms of a mental model and that the process requires some part of the model to be maintained in working memory. (Johnson-Laird, 1983, discusses the differences between a mental model approach and that of Kintsch and van Dijk.) There are indications that reading comprehension skill correlates with the ability to relate linguistic information to a mental model. In an experiment by Merrill, Sperber, & McCauley (1981), less-skilled fifth-grade comprehenders appeared to have more difficulty relating information from sentences to a mental model than better comprehenders did. Nevertheless, the less-skilled comprehenders showed evidence of understanding the words in the sentences, as they would if they had developed only a superficial representation of the meaning (also see Oakhill, 1982).

Since spoken language is more often related to the context in which it is understood than written language is, listeners' mental models are more likely than readers' to be supported by the extralinguistic context. The absence of external support in reading may increase the burden of maintaining currently important information in working memory. However, writing has an important advantage over speech in the provision of linguistic context: The text remains available. Thus, whenever the reader
realizes that previous information is needed, he or she can look
back to recover the content, rather than having to retrieve it
from memory. Print can therefore take over part of the function
of working memory in integration, allowing the reader to recover
antecedent information by retracing through the text. In an eye
movement study by Carpenter and Just (1977), adult readers
performed such regression very precisely, looking directly back
to the place in the text where a potential antecedent occurred.
This indicates that the use of prior text as a support for memory
in the integration of information may be a well-developed ability
in skilled reading.

Implications for Learning to Read

We have examined several differences between written and
spoken discourse which suggest that the ability to develop mental
models from spoken language does not fully or adequately support
the comprehension of written texts. In this section we will
elaborate some potential implications of this suggestion for the
acquisition of reading, focusing again on the roles of the
comprehender's knowledge and the language of the text in the
process of mental model construction. We assume in the following
discussion that similarities between spoken and written discourse
facilitate children's text comprehension, while differences
create areas in which the beginning reader must acquire new
knowledge and skills.
General knowledge and discourse knowledge. We suggested earlier that comprehending different types of texts requires the construction of different types of mental models. We hypothesize that children learning to read bring to the task a background of general knowledge and familiarity with spoken discourse structures that prepares them to construct some kinds of mental models more readily than others.

Consider again the contrast between narrative and expository prose. The intuitive and empirically supported differences in the ease of understanding narratives and expositions may be due to the design of the human mind: Perhaps we are simply better equipped to deal with the kinds of information that narratives convey. Alternatively—or additionally—the knowledge of plans and events that underlies narratives, and the cognitive skill required in constructing mental models to represent them, may be better developed in most people than the knowledge and skills needed to construct mental models of expository prose.

Differences in the availability of the cognitive resources needed to construct mental models for the comprehension of narratives and expositions may be particularly pronounced for young children. Children are likely to be acquainted with narratives and to know something about simple narrative structures because these structures are common in spoken discourse. Moreover, very young children possess implicit knowledge of causation and intention and their roles in real-
Discourse Structure and Mental Models

world events (Gelman & Baillargeon, 1983; Hood & Bloom, 1979; Nelson & Gruendl, 1981), knowledge that is critical for the interpretation of narrative texts. The explicit teaching of simple narrative forms may thus be unnecessary, at least as an adjunct to reading instruction.

The situation for expository prose is different. The infrequency of exposition in spoken language makes such texts unfamiliar to inexperienced readers. The general knowledge of abstract argument structures needed to support the comprehension of expository discourse may be fragile in most young readers, and knowledge of appropriate global text structures nonexistent. Thus, there is little in the way of relevant prior knowledge to support reading comprehension.

Although instructional effort with respect to discourse comprehension may be better centered on expository prose, narratives serve other purposes in the reading curriculum. Stories are a class of narrative designed to entertain, and they are frequently structured to produce enjoyment: The classic mystery story is not written to maximize comprehension but to heighten suspense and curiosity about omitted events. (For an analysis of the structure of stories in terms of the affective states that they evoke in readers, see Brewer & Lichtenstein, 1981, 1982.) The motivational implications of this line of reasoning must be considered when working out instructional programs. It has often been noted (see Gibson & Levin, 1975)
that, if a child is to read large amounts of text and become an independent reader, the material that the child receives should be interesting and entertaining, not just easily comprehended. (See Jose & Brewer, 1984, for research on factors influencing story liking in young children.)

**Discourse cues and local integration processes.** The transfer from spoken to written discourse demands refinement and extension of the ability to integrate information from successive sentences in a text. However, memory deficits similar to those that limit preschoolers' and kindergartners' integration of spoken language also appear in novice readers. Johnson and Smith (1981) asked third and fifth graders to answer questions that required drawing inferences from passages they had read. The third graders were more successful when both premises required for the inferences were in the same paragraph, instead of in different paragraphs. This finding held even when the children were able to recall both premises in response to other questions, indicating that the necessary information was available somewhere in memory. Fifth graders were less influenced by separation of the premises in the text. Johnson and Smith (1981, p. 1221) suggest that older children strategically retrieve previous material, while younger children integrate items of information only when the current one "happens to 'call up' the first from long-term memory or when the two are temporally close and, thus, jointly present in working memory."
Alternatively, older children may be better at maintaining important information from previous text in working memory. If children are more successful in inferring relationships between elements of text and information in the mental model when that information is in working memory, maintenance or reinstatement of important material is crucial in the integration of new information. Adults spend more time reading important than unimportant information in narratives (Cirilo & Foss, 1980; Mandler & Goodman, 1982), and important information is more accessible in memory (Fletcher, 1981). Selecting what is important in order to maintain it actively in memory requires a degree of prior knowledge relevant to the content or structure of the discourse. In types of texts that are less familiar to young readers, such as exposition, active maintenance of central information may be hampered by difficulty in recognizing what is important. Less-experienced readers may thus need help in learning to identify the main points of expository prose and in learning how to maintain them efficiently in memory (for example, through summarization strategies; Brown & Day, 1983).

Sometimes, however, the information needed to interpret and integrate information in text is not maintained and must be recovered from a less active state in memory. When adult readers encounter a reference to previously mentioned information, that information is commonly reinstated in working memory (Chang, 1980; Dell, McKoon, & Ratcliff, 1983; Frederiksen, 1981; Lesgold,
Discourse Structure and Mental Models

Roth, & Curtis, 1979; McKoon & Ratcliff, 1980). If necessary, the antecedent information may be inferred (Clark & Haviland, 1977), and in some cases, the reader looks back to previous material to recover it (Carpenter & Just, 1977). Reinstatement requires engaging in activities appropriate for the identification and retrieval of information in the mental model but not in working memory. Beginning readers may be less likely to engage in such activities, because of the deficits in initiating strategic memory retrieval that are often seen in young children.

Another important component of reinstatement is the explicit or implicit understanding that more information is needed for adequate comprehension. Such understanding includes the ability to recognize that none of the currently accessible information matches the specifications of expressions referring to given information. Some of the difficulties that this may create for young readers can be appreciated by reconsidering the distinction between deictic and anaphoric uses of pronouns.

Deictic uses of pronouns predominate in the speech addressed to young children, where the topics of conversation center on the here-and-now rather than on events displaced in time or space. Karmiloff-Smith (1981) has argued that five- and six-year-old children's use of pronouns is fundamentally deictic, with pronouns taken as pointers to salient elements in the extralinguistic context, not as indicators of coreference with
particular previously mentioned entities, that is, not as anaphors. If the anaphoric function of pronouns is relatively unfamiliar to beginning readers—and Tyler's (1983) work suggests this is a reasonable assumption—they may fail to realize that a pronoun does not refer to any of the information that is readily accessible.

Evidence that lack of skill in processing anaphora can affect text integration in young readers has been reported by Garnham, Oakhill, & Johnson-Laird (1982). Their seven- and eight-year-old subjects fell into two groups matched for age, word recognition, and sight vocabulary but differing in comprehension ability. Each child read one of three versions of a passage that varied in plausibility and coherence. The plausible version described a normal sequence of events involving a young boy playing with a ball, while the other two versions contained an implausible sequence. These implausible versions differed in the degree to which pronouns in the passage could be linked to antecedents within the text. The coherent implausible version was written so that appropriate antecedents for the pronouns could be readily inferred, despite the absence of a normal event sequence, while the incoherent implausible version was written so that it was difficult to locate antecedents. On a subsequent test, the skilled comprehenders did not differ from the less-skilled comprehenders in recall of the basic ideas from the implausible incoherent passage, but they were significantly
better on both coherent passages. The less-skilled comprehenders did no better on the coherent implausible version than they did on the incoherent implausible version, although their performance improved on the plausible passage.

Such findings strongly suggest that younger and less-skilled readers may not deal efficiently with pronominalization in text. Yet a major source of potential guidance in learning to understand anaphora appears to contribute very little to resolving the trouble: Basal reading materials and teaching manuals rarely offer instruction in handling anaphoric relations (Johnson & Barrett, 1981).

Some less-skilled comprehenders may experience enduring problems with the interpretation of anaphoric relationships and other components of the local, sentence-to-sentence integration process. Vipond (1980) found that variability in performance among less-skilled college readers who read and recalled technical passages could be attributed primarily to the difficulty of such local processes (see also Graesser et al., 1980). The ability to integrate information across consecutive sentences in written discourse thus appears to be correlated with reading success.

Conclusion

Our analysis of the problems of discourse comprehension faced by beginning readers has touched on three broad themes that bear a brief summary. These three themes are the general nature
of text comprehension, the discourse comprehension ability of preschool children, and differences between spoken and written discourse.

With respect to the general nature of text comprehension, we argued that the fundamental process is the construction of a mental model that captures the content of the text in a unified representation. Mental models are shaped in part by processes operating on the global structures and local language of texts to integrate new information appropriately. But, equally important, model construction draws on the comprehender's knowledge of such basic categories as human intention, causality, space, time, and logical relations, as well as of such prosaic matters as first bike rides.

With respect to early discourse comprehension, we suggested that, before children learn to read, they are capable of forming mental models from the texts that they hear. However, their ability is limited, perhaps by such factors as restricted general knowledge, inexperience in constructing certain types of mental models, unfamiliarity with various global text structures, difficulty in understanding anaphoric expressions, and deficits in the use of memory.

Finally, differences between written texts and the spoken discourse which children have mastered when they begin school have the potential to exacerbate their comprehension problems. Spoken discourse taps a narrower range of knowledge, it is
structured in different ways, it uses different anaphoric devices, and it provides more contextual support for its interpretation than written texts do. The transfer of discourse competence to reading should be easiest for texts such as narratives that are common in both speech and writing, and that rest on knowledge that young children firmly possess. The major hurdle in the acquisition of discourse comprehension skills may be expository prose. Since most of the knowledge that students are expected to acquire in school is conveyed in that format, facility in dealing with expository discourse represents a crucial step in the development of text understanding.
Discourse Structure and Mental Models

References


Discourse Structure and Mental Models


Discourse Structure and Mental Models


Discourse Structure and Mental Models

48


