HANDBOOK of DISCOURSE ANALYSIS

VOLUME 1: Disciplines of Discourse
VOLUME 2: Dimensions of Discourse
VOLUME 3: Discourse and Dialogue
VOLUME 4: Discourse Analysis in Society

Edited by
TEUN A. VAN DIJK
Department of General Literary Studies
Section of Discourse Studies
University of Amsterdam
Amsterdam, The Netherlands

1985

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
London San Diego New York
Austin Boston Orlando Sydney
Tokyo Toronto
CHAPTER 4

Cognitive Psychology and Text Processing

Gordon H. Bower and Randolph K. Cirilo

INTRODUCTION

The goal of cognitive psychology is to understand how the mind works, how it manages to perform the small miracles of skill we see around us in everyday behavior. Cognitive psychologists take as mysteries what the layman accepts as common abilities—the ability to perceive the world as it is, to remember things, to comprehend social and linguistic events, and to act effectively to satisfy one’s needs. Cognitive psychology is concerned with how people acquire and represent knowledge about their world, how they organize and use that knowledge. The ability to use language in speaking, listening, reading, and writing is one aspect of that knowledge. It is a complex cognitive skill that typically requires several years to acquire and several more to gain facility in.

This chapter focuses primarily on people’s ability to listen to or read coherent texts, to comprehend them, to remember them, and to answer questions about them. In regard to people’s ability to comprehend text, the goal of cognitive psychology is to describe the component processes involved, how they are organized and integrated to support the common feats of text understanding that people perform routinely. Of course, listening to or reading text admits of several different levels of analysis: neurological, phonetic, lexical, sentential, pragmatic. Cognitive psychologists are likely to skip over the neurological or phonetic levels in favor of the lexical or sentential levels of analysis.

Research on language within cognitive psychology falls mainly into three categories or styles. First are empirical studies designed to show that some variable influences language comprehension. An example would be...
be a demonstration that having a picture of a scene available improves the speed with which people can read a description of it with a given level of comprehension and memory. Another would be a demonstration that sentences with pronouns that have ambiguous referents take longer to comprehend.

The second style of research in cognitive psychology is concerned with specifying a small-scale theoretical model of the moment-by-moment operation of a subsystem of a person's language faculty. The subsystem is isolated and studied experimentally to gain understanding of its operation. An example of this would be the studies of Clark and Chase (1972) on how people verify a sentence (like the star isn't above the plus) in comparison to a simple picture (say, of a plus above a star). Experimental psychologists are most comfortable with just this type of situation—a circumscribed theory closely linked to a circumscribed experimental situation—and it provides a display case for their unique contributions. The presumption underlying such activities is that scientific understanding of a complex topic will proceed most rapidly by carving out manageable, small sections of the overall topic for intense investigation. The hope is that the type of theory developed for the subpart will be sufficiently general and consistent in principle with other miniature theories being developed to deal with other linguistic subsystems. The danger is that such miniature theories will become miniature islands isolated from a full scientific account of language competence.

The third style of research in cognitive psychology aims to construct a comprehensive theory of language understanding and text processing. The theory attempts to specify important processes and their organization in sufficient detail that a computer program could be written to simulate all or part of the theory. This characterizes the computer simulation approach followed by Winograd (1972), Schank and Abelson (1977; also Schank & Riesbeck, 1981), Anderson (1976), and Norman and Rumelhart (1975), to mention a few practitioners. Computer models of natural language understanding attempt to specify sufficient mechanisms to perform designated linguistic tasks, relying upon common knowledge and linguistic intuitions to decide when the human task is being accurately simulated by the program. While attracted to such global models of language processing, experimental psychologists are somewhat put off by their apparent complexity, wondering whether the unfathomable human "blackbox" is being replaced by another blackbox (the program) that is just as complex and unfathomable. So, psychologists cull through the programs for basic principles, and then search for ways to collect discriminating experimental evidence for these principles.

In this chapter, we outline the basic cognitive system, its parts and organization, and then sketch briefly a view of how language is understood. We begin with a description of the cognitive system and its operation. We then consider how a single sentence is parsed and understood. Then we take up processing of interrelated sentences in a text.

**THE COGNITIVE SYSTEM AND ITS ARCHITECTURE**

The information processing approach assumes that perception and learning (from linguistic inputs or otherwise) can be analyzed conceptually into a series of stages during which particular mechanisms perform some elementary operation. Given a stimulus input to the system, the operation might be to isolate the figural stimulus from its background, to extract its significant features, to describe it in terms of perceptual primitives, to classify it as an instance of a certain kind, and to associate a meaning with the pattern. What gets passed along from stage to stage is an internal representation of the external stimulus and the context of its occurrence. Theories try to represent the causal flow of events in terms of a flow diagram in which blocks represent component processes, each labeled according to its typical function.

While several flow diagrams have been proposed, a prototypical one is shown in Figure 4.1. The interconnections between components are stipulated by the arrows representing both the flow of control and of information. The components can be divided roughly into a sensory system, a response system, a long-term memory (LTM), and a central processor (short-term memory [STM] and working memory) wherein occurs the active processes we identify with perceiving, memorizing, thinking, and deciding. The diagram is most quickly understood by describing how its parts operate in a simple stimulus-driven task such as reading a series of single unrelated words.

**Pattern Recognition**

The task instructions establish the goal to "read the words presented on a display screen." A plan to achieve that goal is entered into working memory. That plan consists of instructions to look at a certain place (controlling head and eye movements), to name any word presented there, and to continue doing so until told to stop. A word like pencil presented on the display is detected as a change in stimulus energies (a 'stimulus event'); contour-enhancing operations in the retina segregate the visual word as figure from its background. The pattern is laid down
as an icon, or fleeting memory, in a visual sensory buffer, and a set of feature detectors located along the visual pathways begin extracting significant features from the stimulus. These features (vertical lines, angles, intersections) constitute a primitive description of the object.

Next, an identification stage ensues during which the system tries to classify the stimulus. Roughly speaking, pattern recognition is assumed to involve a weighted matching of the current feature list against a likely set of prototypes (idealized patterns) in LTM, with the input being classified according to the name of the best-matching prototype. Accuracy of identification depends on the quality of sensory information extracted and how many alternative prototypes are under consideration. A word will be identified more readily if it is probable and expected in the context. Thus, if the words in the list being read constitute a sensible sentence, then implicit predictions are made and those word prototypes are primed into readiness for matching against the input stimulus. This is why a familiar text, or a text about a familiar topic, is read more quickly. These expectations regarding the next word are being generated on-line from knowledge in LTM.

**Short-Term Memory**

To continue with our naming task, after *pencil* is identified, its internal code becomes active in STM. The plan sitting in working memory is basically a set of rules (technically, sets of condition–action productions) which, when fired, result in the person saying aloud the name of any word code entered into STM. Thus, the articulatory parameters for saying *pencil* are sent to the vocal apparatus, and the word is said. The system then prepares to read the next word and the task continues.

The STM in Figure 4.1 is the active part of the central processor that holds the internal symbols currently in the focus of attention and conscious processing. The STM need not be viewed as a “place” or “register” physically distinct from LTM; STM and LTM may only be two different states (or levels of current activation) of the same memory schemata.

The basic characteristics of STM are as follows: (1) it is the active partition of the memory system, (2) the processor has faster access to the items in STM than to items in LTM, (3) STM tends to preserve the surface perceptual properties of the stimuli as well as their temporal order, and (4) STM has a severely limited capacity, about four to seven symbols or coded items.

The capacity of STM is best characterized in terms of four to seven ‘chunks’ of information. A ‘chunk’ is defined as a stimulus pattern or sequence that the memory system recognizes as a familiar single unit for which a internal code already exists in memory. Thus, *cup* is a single unit to English readers whereas *ucp* is just three letters. The perceptual system appears normally to select the simplest, highest-level, discriminating description of the input sequence. Thus, *doghouse* is chunked as one unit rather than as two words or eight letters. Such units are identified by matching inputs to pattern schemata sitting in LTM.

Verbal items can be maintained in STM, if desired, by focusing attention on them and by rehearsing, covertly going over the items repeatedly. This rehearsal serves to maintain items in STM and to transfer some information about those items to LTM. The information transferred is typically the meaning and the context of occurrence of the items in question. Thus, in a memory span experiment (say, to remember *PZK*), the information transferred to LTM would be equivalent to the proposition ‘List N consists of a token of letter *P* followed by a token of *Z* followed by a token of *K*’.
A critical aspect of rehearsal is that it is under ‘strategic’ control. The person’s goals partly determine what he will select to attend to and rehearse (and consequently remember). Applying these ideas to a reader of text, the reader’s goals would determine which textual information is relevant, important, and should receive special attention, rehearsal, and elaboration. This strategic component of STM may thus explain in part why the relevant, important parts of a text are better remembered.

Working Memory

Figure 4.1 also shows a working memory (or intermediate-term memory), which refers to memory structures that maintain information about the local context, but information that is neither in the focus of active memory nor in the distant edges of LTM. Working memory constructs and maintains an internal model of the immediate environment and events of the past few minutes. This local model serves as a framework or context within which dynamic (small) changes are recorded. The new information updates rather than casts aside the current model.

In regard to text processing, working memory would hold a list of foregrounded topics as well as a list of referents mentioned earlier in the text. Both are needed to look up connections for new expressions and statements. Thus, if a text states John is a teacher, later he can be referred to as the teacher, or as he; but to do so requires the reader to keep track of who is under consideration. The topics list in working memory helps resolve ambiguous expressions (the bank) and definite referents not mentioned previously, as in I’m looking for a restaurant where the waiters wear tuxedos. Many psychological studies have examined the process of reference establishment in comprehension; integral to their interpretation is the concept of working memory.

Also, as indicated earlier in our word-reading task, working memory holds the plan that the person is following in performing some task. The plan is typically a hierarchically structured set of goals, subgoals, and anticipated actions. In reading text, one can have various goals—to comprehend every sentence, or to skim for global gist, or to extract only a specific fact, or to look for typographic or grammatical errors, and so on. For each such goal, people have learned particular plans and routines that are activated in working memory and executed to achieve that goal.

Long-Term Memory

The LTM is conceived to be the repository of our more permanent knowledge and skills. It essentially includes all things that are in memory that are not currently being used. The LTM includes our knowledge of the language, of spatial models of our world, of properties of objects and people, of events of our lives, of perceptual–motor skills, and so on. A major enterprise in cognitive psychology is to indicate the types of knowledge people have, how that knowledge is to be represented, organized, accessed, and used. The two dominant representational formalisms are associative networks and production systems. Let us consider these briefly.

Productions

A production is an IF-THEN rule which states that in case a particular condition arises, then a particular action or series of actions is to be taken. An example would be IF you drive up to a stop sign at a road intersection, THEN brake your car to a stop. The conditions can be any patterns of activity represented in STM such as complex propositions; the actions can be either external or internal moves such as searching a memory location, fetching referents from memory, or activating other concepts into STM. This provides the power to model the flow of thought, problem solving, and plan following.

Each production is similar to an instruction in a computer program. In fact, computer programs can be viewed as a hierarchic sequence of productions. Productions thus provide the operational component in our models of cognitive systems: they form the “motor” that moves the “knowledge engine” through its skilled paces. Most cognitive models of motor and intellectual skills (typing, three-place subtraction) are written as production systems (see, e.g., Anderson, 1981).

Productions are typically learned (similar to stimulus–response habits), are strengthened by successful use, and weakened by nonreinforcement. Productions are stored in LTM; a production is selected and fired when its conditions are matched to the active contents of STM. When several productions are selected at the same time, some resolution or concordant action is required (e.g., choose the stronger one). We will not elaborate further details here; a volume edited by Waterman and Hayes-Roth (1978) illustrates many text-processing models realized as production systems.

Associative Networks

The most popular representation for information in theories of LTM is the associative (or semantic) network. The basic elements of the memory are concepts (nodes, symbols) and relations between concepts (propositions, symbol structures). A concept may be a perceptual primitive, an actional primitive, a primitive logical or semantic relation, or a higher-
order concept built up by relations among these parts. The concepts can
stand for generic terms as well as individual constants. The meaning of
a concept is given partly by the configuration of its relations to other
concepts and partly by the referential conditions necessary for the proper
use and application of the term. In typical representations, concepts are
represented as nodes (or cells in a computer memory) and relations
between concepts as labeled arcs, arrows, or associations between the
nodes.

In such a conceptual network, the learning of a new fact or new concept
is solely a matter of recording a part of its representation in memory,
which is accomplished through a specified configuration of relations among
already-known concepts. An event is represented in memory in terms
of a cluster of propositions describing features of the event. These are
recorded in memory by establishing new associative connections among
new instances (token nodes) of the concepts used in describing the event.

If the concept has many structured parts and several variables, with
constraints on their fillers, then it is usually referred to as a ‘schema’
or ‘frame’. A schema is a structured cluster of generic knowledge about
some object or event. For example, our knowledge about a human face
would be a schema composed of related slots (variable features) to represent
the eyes, ear, nose, and so on. The eye color would be a variable (with
values brown or blue, etc.) as would nose style and mouth size. The
schema would specify prototypical information and correlations among
features. Thus, people with black skin usually have brown eyes, those
with red freckles usually have blue eyes, and so on. In encoding the
face of a new acquaintance, the process is to call forth the generic face
schema and to fill in its slots with the specific characteristics of the
person. This process is called ‘instantiation’, and is the way generic
schemas are employed in recording facts about particular individuals or
events. Some theorists (e.g., Rumelhart, 1977b; Thorndyke, 1978) believe
that text understanding is an example of the reader instantiating generic
schemas regarding the type of text (folk tale, nursery rhyme, detective
mystery, etc.). We will later touch on this schema application view of
text understanding.

Activation

It has already been stated that STM and LTM may only be two different
states of the same memory nodes or schemata. This change in state
occurs through the process of ‘activation’. When the activation level of
sensory elements, concepts, or propositions exceeds some threshold,
they become the contents of consciousness (i.e., move from LTM to
STM). Activation presumably spreads from one concept to another, or
from one proposition to another, by associative linkages between them.
A relevant analogy is an electrical network in which terminals correspond
to concepts or event nodes, connecting wires correspond to associative
relations with more or less strength, and electrical energy corresponds
to activation that is injected into one or more nodes in the network.
Activation of a unit can be accomplished either by presentation of
the corresponding stimulus pattern or by prior activation of an associated
thought.

To illustrate, a simple sentence like An old man smoked a smelly cigar
would be represented graphically (and in LTM) as in Figure 4.2. The
language parser (pattern recognizer) analyzes the sentence into atomic
propositions and records these by subject–predicate links between the
respective concepts. Thus, two entities (individual concepts) are created
in memory, the nodes labeled X and Y in Figure 4.2. Then the atomic
assertions are recorded by linking these new nodes to preexisting concept
nodes by subject–predicate (S–P) (or relation–object) links: X is a man,
X is old, X smoked Y, Y is a cigar, and Y is smelly. Each line in Figure
4.2 is to be conceived of as a new associative link to be learned (perhaps
only briefly established while comprehending). The S–P labels are to be
interpreted semantically as ‘set membership’: the referent of the S node
is a member of the set of referents of the P node. Thus, in Figure 4.2,
X is asserted to be a member of the set of things that smoked Y (smelly
cigars).

Given such a linked structure in memory, it can be used to retrieve
answers to questions by a spreading activation process. Thus, if asked
What did the old man smoke?, activation of the queried concepts in

![Figure 4.2](image-url)
STM is equivalent to direct access to the preexisting concepts of ‘old’ and ‘man’; activation spreads out simultaneously from these two nodes and from the relation ‘smoke’. Two of the activation processes intersect at node X (the concept of the specific old man), so it transmits activation along its link to intersect smoke, placing activation on node Y and thence to the smelly and cigar nodes. When these nodes receive activation above their threshold, the question-answering plan in working memory causes the person to report the names of these concepts as the answer to the question, that is, to say smelly cigar.

Spreading activation is a technique for retrieving associative concepts from memory. It has proven useful in explaining a number of experimental facts about simple question answering (see Collins & Loftus, 1975) and memory retrieval (see Anderson, 1976). Clearly, however, an intelligent processor and editor must be postulated that will discard the unwanted or irrelevant material churned out by an unguided spread of activation through a large, intertwined network.

As indicated, the representation of knowledge, its use in remembering, and the operation of the cognitive system are major topics of theoretical activity in cognitive psychology. Space limitations have prevented us from describing more than a fraction of the ideas in this area. For further reviews of this field, see Anderson (1976, 1980, 1981) and Rumelhart (1977a). We henceforth assume that the reader has enough familiarity with basic concepts to follow discussions in this chapter and other chapters in this Handbook written from the perspective of cognitive psychology. To illustrate this perspective, we take up the issue of linguistic parsing of single, isolated sentences. (Later we deal with sequences of sentences characteristic of text.) We focus on what is occurring in STM as an isolated sentence is parsed.

SINGLE-SENTENCE PROCESSING

We consider now the linguistic parsing and comprehension of a single sentence, either written or spoken. The parser and the lexicon in LTM take as input a linear string of words with surface features (order and intonation) and construct an interpretation and semantic representation of it. We may think of reading as the carrying out of a plan that causes the eyes to rapidly fixate successive words of a sentence. Measures of eye fixations indicate that careful readers fixate nearly all content words of a sentence at least once. A likely hypothesis is that the reader’s eyes remain fixated on a word for as long as it is being processed (see Just & Carpenter, 1980). Thus, gaze durations provide information about on-line processing of single sentences or strings of sentences.

A likely hypothesis about comprehension is that the sensory input (words) is first organized in STM into surface constituents, then underlying propositions are extracted, then referents are looked up, then semantic interpretation of the sentence occurs and is deposited in working memory. Surface constituents are grammatical units like a word, a noun phrase, a verb phrase, a clause, and a sentence, often identified through surface syntactic clues (or intonation in speech). For example, the sentence The old man smoked a smelly cigar has such constituents as man, old man, smelly cigar, smoked a smelly cigar, and so on. Once surface constituents are segregated and their functions identified, the parser tries to coordinate underlying propositions to them. Thus, man is coordinated to the proposition ‘There’s an entity X that is a man’; old man is coordinated to ‘X is a man and X is old’, and so forth. The semantic network representation of such atomic propositions is illustrated in Figure 4.2. Eleven successive constituents are identified and corresponding propositions are extracted, and the parser builds continually in STM a hierarchic representation of the atomic propositions embedded in the larger matrix sentence, as Figure 4.2 illustrates. Once this translation into a semantic medium occurs, the surface string of words is allowed to decay from STM so that only the semantic interpretation is retained for very long.

The steps mentioned above—organizing words into surface constituents, extracting underlying propositions, setting up semantic structures in memory—go on cyclically and in parallel as the reader goes through a text sentence by sentence. The several levels of analysis cooperate and share partial results to help in one another’s task.

Readers and listeners identify constituents by using a combination of syntactic and semantic knowledge. Psycholinguists have adopted different explanatory approaches to parsing. One approach supposes that readers or listeners have a loose-knit set of heuristic rules or parsing strategies that they use as the opportunity arises. An alternative approach tries to implement a complete computational model of parsing such as an augmented transition network. We briefly describe these two approaches.

The strategies approach (see Bever, 1970; Kimball, 1973) says that readers use linguistic clues to control syntactic expectations. Here are several heuristics and types of clues for English syntax (from Clark & Clark, 1977):

1. A function word indicates the beginning of a new constituent.
2. Use affixes to help decide whether a content word is a noun, verb, adjective, or adverb.
3. Use the first word (or major constituent) of a clause to identify the function of that clause in the sentence.
Here are a few semantic strategies:

1. Using content words alone, build propositions that make sense and parse the sentence into constituents accordingly.
2. Identify the verb and look for noun phrases that fit its semantic requirements.
3. Replace known definite noun phrases by their referents as soon as possible.
4. Expect 'given' information to precede 'new' information, unless the sentence is marked otherwise.

There is psychological evidence for the existence of each of these rules. For example, the last rule cited is supported by results showing slower and less accurate comprehension of statements when the order of mention of new versus given information is reversed in a sentence.

Despite the psycholinguistic evidence in its favor, a collection of strategies and rules such as these is incomplete and hardly performs adequately in parsing more than a subset of English or other languages. An alternative theoretical approach searches for a computational model that is both descriptively adequate and sufficient in so far as it can actually parse a significant portion of English sentences. Several computational formalisms have been proposed, including production systems (Anderson, 1976), word- or phrase-based programs (see Riesbeck, 1978; Wilensky & Arens, 1980), and augmented transition networks (ATNs) (see Kaplan, 1973; Woods, 1970). We discuss only the ATN formalism here as it is easiest to describe and has had the widest range of applications.

Augmented Transition Networks

An ATN is a computational algorithm that goes through a sentence left to right, classifying the part of speech of each word, trying to apply syntactic and lexical rules to arrive at the logical relations ('deep structure') being expressed in the surface sentence. The ATN consists of states that are connected by labeled arcs; the analysis of a sentence consists of passing from one state to another as successive words occur. The condition for a given transition is the occurrence of either a certain word (like by in English), or a word of a specified syntactic category (like verb), or a constituent of a specified kind (like noun phrase). To recognize a constituent like a noun phrase or verb phrase may require a call to an appropriate subnetwork, as shown in Figure 4.3. As parsing advances through a sentence, a chart accumulates hypotheses about the probable syntactic function of the words and constituents identified so far. Whenever a given transition occurs in the ATN, certain actions or changes in the developing chart are to be carried out by productions (see the table at the bottom of Figure 4.3). Figure 4.3, from Rumelhart (1977a), illustrates a simple ATN for parsing some active and passive declarative sentences. A more complete parser for English would have many more alternative paths through the network.

To illustrate the operation of this ATN, suppose the sentence to be parsed is The old man was struck by a car. The ATN begins in state S/1 looking for a noun phrase (NP). The NP label on this arc calls the NP subroutine; this finds that the first word the is an article, so arc 8 is taken; then old is found to be an adjective, so arc 10 is taken; then man is a noun, so arc 11 is taken. So far, the productions corresponding to the arcs taken (in the condition-action table below) will cause a structure to be built in working memory, assigning man as the HEAD of a noun phrase, with the and old as a determiner and modifier of that HEAD.

The next word is was, an auxiliary verb, which causes the ATN to take arc 12, to complete the NP subnetwork (thus completing arc 1 and
arriving at state S/2), and to assemble the just-completed noun phrase. The was moves the system to state S/3. The next word, the verb struck moves the ATN along arc 4, causing its production in the table to fire, and to modify the chart by setting the ACTION to the current word (struck) and setting the OBJECT of the sentence to what formerly it believed was the SUBJECT (namely, the first noun phrase, the old man). The next word by causes arc 5 to be followed to S/6. Then a SEEK NP for arc 6 causes the NP subnetwork to be called to handle the car. When this arc 6 is finished, the arc action sets this noun phrase to be the logical subject (ACTOR) of the sentence. Arriving at S/4, the ATN notices there are no more words to be processed in the sentence, so arc 7 is taken, which assembles the logical deep structure, namely, ACTOR (the car), ACTION (struck), and OBJECT (the old man). Since the sentence is at an end and the ATN has moved from the initial state to its terminal state, we can say that the ATN has accepted or recognized this word string as a grammatical sentence according to its rules. If the sentence had been ungrammatical, then the ATN would have met some obstacles in the analysis and been unable to continue to the end.

ATN grammars are modular and can be built up piecemeal by adding more subroutines to analyze other grammatical constructions (relative clauses, complements, questions, etc.). Although no ATN grammar has ever been built capable of analyzing all of the English language, systems like that illustrated in Figure 4.3 have been expanded and elaborated to deal with a significant domain of English (see Woods, Kaplan, & Nash-Webber, 1972). From a psycholinguist’s perspective, an appealing feature of an ATN is the natural way the grammar of the language fits together with the rules for processing sentences. A second attractive feature of an ATN is that it gives a natural way to characterize readers’ syntactic expectations about what is coming up next in a sentence. The arcs leaving a given state may be rank-ordered in terms of their expected uses; this forms the basis for syntactic heuristics. These differing expectations are revealed in many psycholinguistic experiments. Thus, in the soldiers man the barricades, the word man could be interpreted as a noun or verb, but after soldiers the ATN expects a verb, and so man is so interpreted. Such heuristics are misled by “garden path” sentences like The old man the barricades because old is interpreted first as an adjective and man as a noun (see the arcs of Figure 4.3), and the ATN halts when it comes to the barricades because it is looking for a verb. It then backs up, reinterprets old and man as a noun and verb, respectively, and continues with the new parse. Psychological experiments demonstrate that people do halt their reading at the points predicted by an ATN, and return to

that point in the sentence where they went astray. Eye regression records reveal this process clearly.

Much psycholinguistic data about the difficulty of parsing is explained by an ATN. For example, an ATN explains why a right-branching construction like (1a) is much easier to understand than a center-embedded sentence like (1b).

(1) a. The cat chased the mouse that lived in the house that Jack built.
   b. The house the mouse the cat chased lived in was built by Jack.

The center-embedded sentence requires the ATN to repeatedly call and interrupt the noun-phrase subroutine (see Figure 4.3) and then to pair off appropriately the stack of noun phrases with a stack of verb phrases. Further evidence for the ATN expectations is that when asked to continue a sentence fragment, people will fill in a word of the grammatical category predicted by the ordering of arcs in the ATN (see Stevens & Rumelhart, 1975). Similarly, when reading aloud, the errors people commit are invariably substitution of a word falling within the syntactic category that the ATN predicts at that point in the sentence.

Sentence parsing is an active research field, and current theories have led to many refinements of ATNs and beyond. There is also an active debate over how accurate ATNs are in accounting for psycholinguistic data (see Fodor & Frazier, 1980; Frazier & Fodor, 1978; Wanner, 1980). For our purposes here, the details of the parsing mechanism are not so important as is the form of its output. As noted before, the output of the parser is a cluster of propositions in working memory; these are represented in terms of labeled associations among the concepts (see Figure 4.2). This associative structure may be transferred to LTM; that is, the temporary links established during comprehension may become permanent. After a sentence is processed and its graph structure has been set up in working memory, its surface form in STM is permitted to fade. This fading occurs rapidly if attention is turned to a following sentence. In text, with streams of related sentences, the parser has the job of connecting successive sentences. We now turn to discussion of the processes that connect successive sentences in coherent text.

**INTERPROPOSITIONAL CONNECTIONS**

The view of comprehension presented thus far is in no sense unique to discourse. That is, most accounts of single-sentence parsing are just
that: theories of the comprehension of single sentences isolated from any larger context. However, it should be clear that the meaning of a discourse cannot be adequately represented by a sequence of independent propositions (see de Beaugrande, this Handbook, Vol. 1). At the very least, a comprehender must recognize that the propositions are interconnected. In this section, an overview of the cognitive processes assumed to make interpropositional connections is given. Evidence for this picture is also provided.

Many sorts of interpropositional connections exist: referential, temporal, causal, and so on. Within psychology, however, most research has focused on the establishment of referential connections, perhaps because the to-be-connected elements are relatively obvious in discourse and, hence, amenable to a rather straightforward theoretical treatment.

Reference Establishment

As new propositions are processed, they must be integrated with earlier information in the text or with other knowledge present in LTM. At the referential level, this means that the comprehender must first determine whether a referent is given (information encountered earlier in the text) and, hence, already represented in working memory or new and, hence, not present in working memory (Clark & Haviland, 1977; Haviland & Clark, 1974). Cues in the text will often signal what is given and what is new (e.g., definite versus indefinite articles).

If the referent is taken to be new, then it cannot be linked to anything previously encoded from the text. Instead, a defining proposition (see Kieras, 1981) must be constructed for it, in essence, a new node in memory similar to ‘X is a man’ or ‘Y is a cigar’. If the referent is given, however, such propositions are already present and the comprehender must search for the appropriate place to attach the new information. Just and Carpenter (1980) suggest that such searches are guided by several strategies. Comprehenders may first look for connections with antecedents that have been repeatedly referred to, are recent, or are topical. In fact, Kieras (1981) suggests that a topics list is maintained in STM for just such purposes. Consequently, the search for coreferents need not be a simple, linear one.

Moreover, the search will be constrained by the limits of the cognitive system (see Lesgold, Roth, & Curtis, 1979). That is, STM is limited in capacity. Hence, only a subset of the already-processed propositions will be active at any one time. If a referential connection can be made to an active proposition, the search will be relatively brief, leading to an immediate match. However, it may be that referential connections must be made to propositions that are no longer active. In such a case, the earlier memory structure must be reinstated in STM before an identification can be made. This is a reinstatement match. It relies upon the retrieval processes discussed earlier. Still, memory searches may fail. It may be that the cue to “given-ness” is inappropriate or the search cannot locate the needed memory structure. In such cases, inferential processes must be invoked to make the connection. This corresponds to Haviland and Clark’s (1974) notion of a ‘bridging inference’.

The evidence suggests that referential connections are made as soon as new referents are encountered. However, comprehenders occasionally make assignments that prove wrong on the basis of later information. Consequently, Just and Carpenter (1980) have held that a running representation of each clause is maintained with updates as each new word is processed. They also postulate that special processing occurs at the ends of sentences. Any gaps or inconsistencies that could not be resolved within the sentence are handled at this point.

This is only a rough outline of referential processing. In the following paragraphs, however, we consider some of the research that has been generated by this simple picture.

New Referents

First, consider the case of propositions whose referents are all new. As stated earlier, such propositions cannot be simply integrated with already processed material. Instead, new nodes must be created for the new referents. If this is so, then propositions containing new concepts should take more time to process than those containing previously mentioned concepts. A study by Kintsch, Kozminsky, Streby, McKoon, and Keenan (1975) supports this notion. They compared the reading times for passages with referential repetitions and pronominalizations to those for passages of the same length without such repetitions. As predicted, the passages with repetitions took less time to read. Kiersas (1978) also found longer reading times for sentences containing only new referents. In short, when referents do not corefer with earlier concepts, extra processing is required. This is consistent with the creation of new nodes in memory.

Prior Coreferents

Next, consider the case of propositions that are coreferential with earlier propositions. As noted, referential ties can be drawn through immediate matches, if the propositions containing the critical referents are both in STM, or through reinstatement matches, if the earlier referent
must be retrieved from working memory. Initially, it can be asked whether there is evidence that referential connections are being drawn during comprehension. Carpenter and Just (1977) have demonstrated this by examining eye fixations during reading. They assume that eye fixations can be external indicators of the internal integrative processes. If so, referential ties should be indicated by regressive eye fixations between the current referent and its earlier occurrence. Of course, this assumes the current referent unambiguously points to a single, earlier referent. If not, regressions should occur to any referent that could be appropriate. Just and Carpenter compared regressions from referents with singular and ambiguous coreferents. They found that regressions were more determinate in the former case than in the latter. Thus, regressions are obtained between coreferents, regressions that increase in specificity as the linguistic context dictates.

Immediate versus Reinstated Referents

Now consider the distinction between immediate and reinstatement matches. These two processes are differentiated by the activation state of relevant segments of text with respect to STM. Theoretically, one can influence this state by varying the distance between the current sentence and the earlier sentence containing the coreference of interest. The farther apart they are, the more likely it is that the earlier sentence has become inactive and, hence, will require a reinstatement match. Clark and Sengul (1979) had subjects read a three-sentence context paragraph, press a button, read a target sentence, and then press a button when they had understood the target. They varied whether a referent in the target sentence was mentioned earlier in the first, second, or third sentence of the context paragraph and measured the time to understand the target (the time between button presses). They found comprehension time to be fastest when the referent was mentioned in the third sentence and slowest when in the first, comprehension time for a referent in the second sentence falling in between the two. Similarly, Carpenter and Just (1977) found that the time to decide if a sentence was consistent or contradictory with preceding information in a text increased with the number of intervening sentences between the target and an earlier sentence to which it had to be linked referentially. Thus, the more likely it is that referents must be linked on the basis of a reinstatement match, the longer it will take to understand the current sentence.

But is reinstatement actually taking place? Work by McKoon and Ratcliff (1980b) demonstrates this rather elegantly. They used an activation (or priming) procedure in which the subject read a paragraph sentence by sentence and was then presented with a single test word from the first sentence of the paragraph for recognition. They varied whether the last sentence of the paragraph referred anaphorically to the first sentence. If it corefers and there is reinstatement, then the earlier referent should be activated in memory and, hence, should be easier to recognize. This is what they found. Recognition time for a referent in the first sentence was faster when the last sentence had referred back to it than when it had not. Furthermore, the entire proposition was reinstated and not just the referent, for recognition time was also shortened for words in the same proposition as the referent of the anaphor in the last sentence. Thus, there is evidence for the occurrence of reinstatements.

As stated earlier, though, memory searches may fail. For example, when cues to "given-ness" are inappropriate, memory searches should occur even though there is no antecedent to be found. Haviland and Clark (1974) created sentence sequences in which definite references (which should be taken as given information) either did or did not have explicit antecedents. They found that the lack of antecedents resulted in longer comprehension times for the sentences containing the definite references. For example, The beer was warm (the definite article indicating "given-ness") takes longer to comprehend following John got the picnic supplies out of the trunk than following John got the beer out of the trunk. In the latter case, the connection between the two sentences is clearly defined by the recurrence of the beer. In the former case, however, there is no direct antecedent for the beer, rather a connection can only be made by inferring a relation between picnic supplies and beer. Hence, inappropriate signals to "given-ness" result in additional processing, the invocation of inferential processes to make the indicated connection.

There is also empirical support for special integrative processing at the ends of sentences. For example, when subjects read texts word by word or phrase by phrase, they tend to pause longer at the word or phrase ending the sentence (Aaronson & Scarborough, 1976; Mitchell & Green, 1978). In sum, there is ample evidence for the processing account given above (see also Cirilo, 1981; Vipond, 1980).

Referential Coherence and Integration

Referential coherence should be reflected in memory representations for texts. To use Kintsch and van Dijk’s (1978) terminology, the product of the indicated processing is a coherence graph, a network of propositions interconnected referentially. Typically, coherence graphs are represented hierarchically. Propositions to which many others are connected occupy the higher levels of the graph.
If referential coherence is important in textual representations, information should be better integrated where coreference exists than where it does not. In an early test of this, de Villiers (1974) had subjects read a list of sentences that could form a referentially coherent passage but did not tell them this. For some subjects, the sentences contained definite articles, while for others they contained indefinite articles. Since definite articles should evoke attempts at referential integration, a coherence graph is more likely to be constructed in the former case than in the latter. The results supported this hypothesis. With definite articles, recall for the sentences was better and subjects tended to make more intersentence lexical substitutions (interchanging coreferents expressed by different words). In short, when readers have reason to believe that referential ties can be made, they will tend to do so, thus producing a greater degree of integration of the information.

This notion is further supported by Lesgold (1972), who studied pronominalization as a connective device. He compared the integration in memory for sentences whose propositions were or were not related by pronominal reference. Thus, *The postman whistled a tune and he (or George) was tired* links *tired* with *postman* in the *he* sentence but not the *George* sentence. Lesgold found that pronominal reference resulted in integrated representations (an equal likelihood of recall of lexical items in the same versus other underlying sentence propositions), while the lack of it resulted in incomplete integration.

**Distance in a Coherence Graph**

Although such research emphasizes the importance of coreference to memory for texts, it does not demonstrate that coherence graphs are an apt depiction of textual representations. For example, coherence graphs define different relative distances among propositions. Two propositions that corefer will be closer than two that do not. This should be true even when coreferring propositions are separated by several sentences in the surface text. Indeed, two propositions that are relatively close in the surface text may end up relatively distant in the coherence graph. Nonetheless, their closeness in the underlying coherence graph should indicate their psychological distance.

McKoon & Ratcliff (1980b) tested this distance implication using an activation (or priming) technique similar to that mentioned earlier. Specifically, the closer two propositions are in the coherence graph, the more the appearance of an item in one should prime or facilitate the subsequent recognition of an item from the other. Subjects read two unrelated paragraphs and then were tested on single words presented one at a time. Subjects had to decide whether or not each word had appeared in either of the paragraphs. However, the structure of the paragraphs was varied. Some had a linear structure (see Table 4.1). Thus, the amount of priming between two words appearing contiguously in the test list should depend only on the linear distance between them (e.g., *Tom* should be a better prime for *boy* than *flower*). Other paragraphs had a structure in which certain nouns were equally far apart in the surface text as other nouns but closer in the coherence graph (see Table 4.1). A noun close in the coherence graph to another noun should be a better prime for that noun than a noun that is equally distant at the surface level but more distant in the coherence graph (e.g., *nurse* should prime *cabinet* better than *patient* primes *ground*). The data were consistent with these hypotheses. In a subsequent study using the same procedure, McKoon and Ratcliff (1980a) also demonstrated that words from referentially connected propositions in LTM primed one another better than words from non-coreferential propositions in the same paragraph. In short, coherence graphs bear some relation to textual representations in memory.

**Table 4.1**

Examples of Two Different Text Coherence Graphs

<table>
<thead>
<tr>
<th>Linear structure</th>
<th>Branching referential structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 <em>Tom saw a boy.</em></td>
<td>P1 <em>The doctor called a nurse.</em></td>
</tr>
<tr>
<td>P2 <em>The boy threw a ball.</em></td>
<td>P2 <em>The nurse pushed a wheelchair.</em></td>
</tr>
<tr>
<td>P3 <em>The ball hit a dog.</em></td>
<td>P3 <em>The wheelchair held a patient.</em></td>
</tr>
<tr>
<td>P4 <em>The dog bit a girl.</em></td>
<td>P4 <em>The patient had a cut.</em></td>
</tr>
<tr>
<td>P5 <em>The girl dropped a doll.</em></td>
<td>P5 <em>The doctor opened a cabinet.</em></td>
</tr>
<tr>
<td>P6 <em>The doll smashed a flower.</em></td>
<td>P6 <em>The cabinet contained a bottle.</em></td>
</tr>
<tr>
<td></td>
<td>P7 <em>The bottle struck the ground.</em></td>
</tr>
</tbody>
</table>

*Adapted From McKoon and Ratcliff, 1980b.*
To summarize, referential links tend to be drawn when cues in the text indicate it. When such links exist, the result is a better integrated representation of the text. Furthermore, when referential processing should be more time consuming, evidence for such difficulties can be found.

As stated earlier, however, coreference is only one form of textual coherence relevant to comprehension. Unfortunately, relatively little psychological research has been done with connective devices in text other than coreference. For example, Haberlandt and Bingham (1978) have established the importance of verbs to coherence, while Black and Bern (1981) have studied the role of causal connections between sentences. Clearly, much more needs to be accounted for. Still, the ideas described above may provide a foundation on which to build more comprehensive accounts of text processing.

Macroprocessing

Another aspect of processing to be considered is macroprocessing (Kintsch & van Dijk, 1978), since it is often assumed to operate on coherence graphs. It has been hypothesized that, during comprehension, special processes reduce all the propositions contained in a coherence graph to a smaller set of high-level propositions or macropropositions. This set of propositions (which can be defined at several levels of abstraction) would describe the global organization of the text, its gist, thus making summarization and long-term memory of the ideas in the text manageable tasks.

To derive macropropositions, processes must be postulated that can abstract and generalize from the detailed text information to its more embracing concepts. These macroprocesses must occur simultaneously with the low-level processing. Indeed, to a certain extent, the two types of processing must be highly similar. That is, sequences of macropropositions must be coherent with one another. Hence, comprehenders must make referential connections among them. In particular, memory searches, reinstatements, and required inferences will influence processing at this level as well. Vipond (1980) has evidence supporting these ideas.

As to the actual creation of macropropositions, it is assumed that propositions are first differentiated according to their relevance to the topic of the text. To do this, comprehenders must use expectations about what sorts of things are relevant to the topic. Occasionally this is a matter of filling in the slots (values) of a schema or frame to deal with a particular type of text. Thus, newspaper stories about automobile accidents may call up a schema with variables to be filled with time and location of the accident, participants, its cause, extent of damage, and disposition of the victims. Once relevance assignments have been made, irrelevant propositions are deleted (or deactivated) while relevant propositions are incorporated into the macrostructure, perhaps after being modified by other macroprocesses, thus forming a macroproposition. The resultant macrostructures are hierarchical since macroprocesses can be reapplied. One of the macroprocesses is generalization over irrelevant details that are not preconditions for later understanding. This may involve deletion of qualifiers and of specific names, dates, places, or replacing constants by their superordinate categories (e.g., replace apples by fruits).

This aspect of comprehension is dealt with more thoroughly elsewhere (Kintsch, this Handbook, Vol. 2). Still, it should be noted that macroprocessing has not been studied as extensively as referential processing, for example. It is far more difficult to give a detailed process model that will yield macrostructures than one that will yield referential structures. As has been implied above, though, high-level knowledge structures, schemata, are crucial to the differentiation of relevant from irrelevant propositions. The nature and use of such schemata in text comprehension and memory are the subject of the next section.

GLOBAL TEXT STRUCTURES

A schema is a data structure for representing the generic concepts stored in memory (Rumelhart, 1980). Although schemata represent knowledge at many levels of abstraction, research in text comprehension has focused on schemata dealing with relatively high-level aspects of text structure. In this section, these schemata and their use is briefly described. Subsequently, some of the work bearing on these ideas is outlined.

Story Schemata

The schema familiar to most researchers in discourse is that assumed to describe stories (see Gulich & Quasthoff, this Handbook, Vol. 2, for a fuller treatment). This schema depicts the prototypical story and variations on it. It identifies the units that constitute stories, the sequencing of these units, and the types of connections that will typically occur between units. It is assumed to develop through experience with stories and the types of real-world event sequences found in stories.

Story grammars have been widely used to represent story schemata (Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1979; Thorn-dyke, 1977). In story grammars, information about story structure is
represented as a set of rewrite rules, essentially a set of rules for segmenting, identifying, and manipulating story constituents. Although this format is important, its use should not be taken as a claim that the story schema in the comprehender’s head is a story grammar. Typically, one treats the formalism as though it specified the parts of the schema and their interrelationships in order to see if anything can be learned as a result (Bower, 1976). It provides a method of study rather than a complete theory. Consequently, those who argue against the formalism (e.g., Black & Wilensky, 1979) should try to indicate a more fruitful alternative for study.

The most common alternative to story grammars is the event schema (e.g., Lichtenstein & Brewer, 1980). Earlier we stated that the story schema must develop in part through experience with real-world event sequences. Consequently, it can be argued that story understanding is only a special case of event understanding. The claim would be that people understand story structures using their general knowledge about planning and problem solving found in event sequences (i.e., knowledge about how actions are related to one another to organize behavior) rather than by use of text-specific knowledge. Lichtenstein and Brewer support this view by demonstrating that memories for simple event sequences presented on film show properties very similar to memories for stories. Some would argue, then, that stories are no different from descriptions of other event sequences except for their evocation of particular emotional responses like interest, suspense, and surprise (Morgan & Sellner, 1980). Whatever the case, stories are assumed to be understood by way of the typical information patterns represented by schemata.

Expository Schemata

If this is true for stories, it would not be surprising that the same holds for expository prose. That is, there should be schemata which are used in comprehending different types of expository texts. In general, such schemata are hypothesized to be organized around prototypical patterns of argumentation (see Kopperschmidt, the Handbook, Vol. 2). However, the only theory of expository structures that has yet had any impact in psychology is that of Grimes (1975) and Meyer (1975). In this theory, the higher levels of text structure are characterized in terms of rhetorical predicates. These predicates relate textual information by specifying superordinate principles to which the rest of the text is made subordinate. Examples of top-level forms are problem–solution, question–answer, attribute-listing, and contrasting alternatives.

For example, the problem–solution rhetorical form defines two high-level ideas in a text: the statement of a problem and then the problem’s solution. Hence, any crucial element of such a text should be either part of the problem statement or part of the solution.

In our account, these top-level structures are schemata which authors use to organize their texts. At comprehension, readers will use similar schemata to determine the text's global message and how this message organizes the rest of the content. In short, the understanding of exposition is no different from the understanding of stories except for use of schemata specific to such texts. Research finds that good readers who score high on comprehension also organize their recall using the rhetorical structure of the passage.

Schema in Comprehension

But how are schemata used in comprehension? According to Rumelhart (1977b), schematic comprehension consists of selecting schemata to account for the text to be understood and verifying the appropriateness of those schemata. Consider the case of a story. As its first sentences are read, they should activate certain schemata which in turn suggest that other, high-level schemata are appropriate for the text at hand. For example, the occurrence of Once upon a time . . . might activate a schema for the setting of a fairy tale. Consequently, the fairy tale schema should also be activated since the setting schema is one of its components. The activation of a high-level structure on the basis of something in the text itself (i.e., data-driven processing discussed by Bobrow & Norman, 1975) corresponds to schema selection. In essence, an initial hypothesis is made about which schema can account for the text.

Once the high-level schema has been activated, it will activate its subschema. The assumption is that these subschemata will be able to account for some of the input text. This activation from the top can be viewed as the creation of expectations about what will be found in the text (so-called conceptually-driven processing). In the case of stories, one would expect a series of episodes describing a protagonist’s attempts to achieve some goal by overcoming obstacles. Of course, the activation of the episode schema results in the activation of its components, and so on.

The remainder of comprehension can be viewed as an interaction between these expectations and the input text. This is the substantiation procedure. That is, the expectations must be confirmed in the text (e.g., there must be some segment of the text that can be understood as an episode). With sufficient confirmatory evidence, the interpretation offered by the schema is substantiated. If enough disconfirmatory evidence is collected, however, the current schema is deactivated and a more promising alternative is found.
The product of comprehension will consist of an instantiation (possibly fragmentary) of the schema or schemata used during processing; this is essentially a copy of the schema with its variable slots filled by segments of the input text, with nonfitting elements noted. Retrieving such a representation would consist of finding these stored traces in memory and using the available schemata to reconstruct the original interpretation and/or the original text. That is, the traces must be found; then it must be determined what schema was used to encode them; finally, an attempt is made to fit the traces into the schema’s component slots.

Whatever the plausibility of this specific theoretical account, it is apparent that schema processing will not always be successful. Texts will fail to be understood for many reasons; major reasons for failure are that they provide insufficient cues to the appropriate schemata or the comprehender does not possess the crucial schemata. Many studies with younger children (e.g., Kernan, 1977; McClure, Mason, & Barnitz, 1979; Poulsen, Kintsch, Kintsch, & Premack, 1979) show that their textual schemata are more poorly developed and used than those of older children. Clearly, we must be concerned not only with how schemata are used but also how they are formed. Rumelhart (1980) has offered several hypotheses as to how schemata develop.

To summarize, it has been argued that comprehension consists of building a schematic representation for a text. Later, if the text must be remembered, the representation and schematic knowledge will be activated and used to guide retrieval. This is a very rough picture, but a large body of research supports it. In the following, three lines of evidence are considered: studies dealing with schematically well-formed texts, studies dealing with schematically malformed texts, and studies dealing with schematic elaborations on texts.

Research Supporting Schema Theory

Well-Formed Texts

First, consider texts that conform to schematic expectations. By definition, such texts are well formed. Consequently, the schema should be maximally effective during comprehension. Moreover, the structural analysis provided by the schema should roughly describe the representation of a comprehended text in memory. If so, the properties of remembered texts will correspond to those of the schema. Much research has been motivated by the search for such correspondences.

Typically, the structural descriptions inherent in textual schemata are hierarchic. That is, some propositions are superordinate to or more important than others. If texts are processed in terms of such hierarchic structures, then this fact should be reflected in the reader’s memory for them. Specifically, it has been argued (Bower, 1976; Thorndyke & Yekovich, 1980) that recall will start at the top of the hierarchy and work its way down by associative cuing. Retrieval paths will be longer to the low-level propositions. Hence, their retrieval probability will be lower. This has proven to be the most consistent finding in the literature. Whether a narrative schema (Rumelhart, 1977b; Thorndyke, 1977) or an expository schema (Meyer, 1975) is applicable, high-level ideas are recalled better than low-level ideas. Thus, the result of comprehension reflects the hierarchic nature of the structural descriptions.

Schemata also postulate constituent structures for texts. That is, well-formed texts are divisible into smaller constituents. Therefore, comprehenders should be sensitive to these units and encode texts in terms of them. The episodes found in narratives have been the most studied of these units.

For example, Glenn (1978) independently varied a story’s length, as defined by its number of simple sentences, and the story’s episodic structure, as defined by the types of information in the sentences and the intersentential relationships. She found that variations in length influenced the number of statements recalled but not the organization of those statements. On the other hand, variations in structure influenced the organization of information in recall but not the amount recalled. Similarly, Black and Bower (1979) have found that the recall of an episode’s actions depends on the length of that episode but not on the length of other episodes in the story. Hence, story statements cluster into separate chunks (corresponding to episodes) in memory.

But even if hierarchic and episodic structures are evident at recall, it is not clear what the locus of the effects is. The schematic influence could be present at encoding, retrieval, or both. Recent studies, however, have indicated that analogous effects occur during encoding when retrieval is not an issue.

Structural Influences on Encoding

First, the hierarchic structure of a text is reflected in its immediate processing. Cirilo and Foss (1980) compared the reading time for a sentence when it occupied a high-level position in one story to the reading time for the same sentence when it occupied a low-level position in another story. They found that the sentence took longer to read when it occupied a high-level (important) position in the story structure. Just and Carpenter (1980) obtained a similar result with expository prose. Perhaps compre-
hinders spend less time on low-level propositions because they recognize them as irrelevant. Alternatively, high-level propositions may take longer to integrate into the accruing context than low-level propositions. Being globally relevant, perhaps they generate more connections to other parts of the text. Drawing more connections will take more time. Whatever the case, the hierarchic property of schematic representations is found both at encoding and retrieval.

The comprehension of episodic structures has also been investigated directly. Haberlandt (1980) measured the reading times for sentences constituting simple, two-episode stories. After subtracting extraneous contributions to reading time like the number of words in a sentence, he found that reading times were longer at the boundary nodes of an episode than would otherwise be expected. This was attributed to greater cognitive work in “wrapping up” or encoding the boundary nodes. Specifically, at the beginnings of episodes, there are often changes in perspective, characters, and the like, the establishment of which require extra cognitive resources. At the ends of episodes, the comprehender will attempt to organize the sentences into a single, high-level node, summarizing or wrapping up the gist of the episode. Again, the processing load will be greater at this point. In sum, the episodic structure of a story influences the course of comprehension and does not simply organize recall.

Processing Malformed Texts

So long as texts are schematically well formed, schema theory predicts they will be readily comprehended, with most readers agreeing on the relative importance of different sentences. However, suppose that a to-be-comprehended text violates the conventions of a textual schema. Such malformed texts should be difficult, if not impossible, to understand. In order to test this, the applicability of the appropriate schemata must be manipulated. The text’s structure must be completely lacking or sufficiently distorted to prevent easy recognition. A number of techniques have been used to accomplish these aims.

One method of eliminating structure is to arrange the constituents of a well-formed story such that their order is nonschematic. This can be done at a number of levels. Thorndyke (1977) randomly permuted the sentences constituting stories. Stein and Nezworski (1978) created stories in which one story category occurred out of its ideal location. Mandler (1978) used pairs of two-episode stories in which during presentation the nodes constituting each episode were presented in interleaved fashion rather than separately. All of these studies found that comprehension and recall for the ideas in malformed stories was poorer than that for the same ideas in well-formed stories. Moreover, the greater the deviation from the ideal, the greater the decrease in recall accuracy. In short, when the relevance of a textual schema is not readily apparent, memory will be relatively poor.

Despite such problems, comprehenders still attempt to apply textual schemata to malformed texts. This is clear when the recall protocols are examined more closely. In particular, it is found that subjects tend to insert missing elements and reorder the story statements in recall to conform to their ideal order. In other words, schematic considerations influence how information is output such that the distortions at input are lost or corrected. However, this tendency decreases as schematic violations increase. Apparently, large deviations are not corrected, perhaps because the reader remembers that the text was substantially malformed.

A second method of eliminating schemacity relies on the cultural specificity of textual schemata. Other cultures may have very different conventions underlying their texts. When a text does not conform to the schema of one’s own culture, comprehension should be poor. Accordingly, Kintsch and Greene (1978) found that for American college students the quality and intersubject agreement of summaries for stories corresponding to familiar schemata of European culture were better than those for Apache Indian folktales whose connections seem more arbitrary and less predictable. However, one must be careful in interpreting cross-cultural comparisons. Since Apache subjects were not tested, one could conclude simply that the European stories were more comprehensible than the Apache folktales to any readers regardless of their culture. Also, as Johnson and Mandler (1980) have noted, the Apache schema may be the same as ours but the interpretation of the content relies on other culture-specific knowledge which we do not have. Thus, there is as yet no convincing evidence that textual schemata differ fundamentally across cultures.

How do deviations from expected structures create difficulties? Violations upset expectations and cause comprehenders to spend more time explaining or resolving the unexpected elements than processing the rest of the story. However, the memory difficulty could also lie in retrieval. The parts of a well-formed text could act as cues for the retrieval of later parts. Such cues would not exist for texts which are structurally incoherent.

There is some evidence for the former account. Kintsch, Mandel, and Kozminsky (1977) scrambled the paragraphs constituting stories. They found that, when given free reading time, subjects summarized scrambled stories just as well as well-formed stories. However, scrambled stories took longer to read, suggesting that time was needed to resolve the
confusions. Briefly, violations of schematic expectations create difficulties that are not seen when all the expectations are met.

Schema-Directed Elaboration

A final important aspect of schemata is that they contain information that not every text will explicitly realize. Consequently, comprehenders may elaborate on the presented material, generating plausible information consistent with the current schema but not necessarily true. Past experience is used to understand a text and as a result more information is added to its memory representation. For example, a comprehender may infer character motivations, embellish personality descriptions, and so on. The amount and type of elaboration will vary from person to person depending on his or her prior experience with related contexts. If the situation is highly regular (e.g., scriptal situations as described by Schank & Abelson, 1977), however, the nature of the elaboration will be highly similar across readers. According to Reder (1980), readers use such elaborations to find connections among sentences, generate expectations about later input, and aid in retention by creating a richer representation.

One common method of testing this view directs subjects to make certain elaborations (i.e., use certain schemata) rather than others. The expectations are that memory will be better for material consistent with the preferred elaboration and that inferences will be added to the representation on the basis of the preferred elaboration. Bower (1978), for example, reports an experiment in which subjects were given stories consisting of episodes. Half the subjects were given prior information that would suggest certain schemata were relevant, such as that the main character (a college coed) had recently discovered she was pregnant. The meaning of some of the episodes (e.g., going to the doctor) could be very different if the heroine is thought to be pregnant.

As predicted, subjects given the prior information produced more inferences at recall consistent with the pregnancy theme. Furthermore, they recalled more of the episodes related to the theme. This is to be expected if the prior information called up a schema used to make elaborations during comprehension. Similarly, manipulations of the reader’s perspective or bias have also been shown to influence comprehension and memory (e.g., Anderson, Reynolds, Schallert, & Goetz, 1977; Pichot & Anderson, 1977; Schallert, 1976).

In sum, it is apparent that high-level knowledge structures are used in understanding and retrieving texts. Texts consistent with these structures are well remembered and easy to understand. Conversely, texts inconsistent with these structures are poorly understood and remembered, even creating a tendency for subjects to distort their memories toward more schematic forms. Finally, information in texts may be elaborated upon and remembered in terms of the schemata used at comprehension. Clearly, text understanding is as much a product of the knowledge structures of the reader as it is a product of the text itself.

CONCLUSION

In this chapter, we have given an overview of text processing as cognitive psychologists understand it. We have focused upon the architecture of the cognitive system, the cognitive processes that mediate comprehension, and the evidence gathered to test these views. In so doing, we have smoothed over the differences that exist among individual theorists. Some may choose to examine different aspects of the same problem. Others may examine the same aspects but in very different ways. We have attempted to capture the most commonly held picture of text processing.

However, it must be remembered that most of the psychological research in this area is of very recent origin. Consequently, the theoretical ideas must be regarded as preliminary and subject to refinement. In time, we may develop very different notions of the cognitive processes that operate on texts, the nature of textual representations, and so on. Despite this limitation, the empirical data generated thus far will always be relevant. Any new theory must be able to encompass results obtained within the frameworks of old theories.

Where, then, are the new theories coming from? Perhaps the most promising work is being motivated by the concern that scholars dealing with the same problem across disciplines should be interacting with one another. Although cognitive psychologists have drawn heavily from other disciplines in the past (e.g., ATNs from computer science and notions about textual structures from linguistics), only recently have more formal lines of communication been established. This is evidenced by the growth of the interdisciplinary field of cognitive science. It seems reasonable that this pooling of resources should significantly accelerate progress in the field. At this point, however, it is probably premature to point out any particular new approach.

Modern civilization demands that a person acquire and use language. It is imperative that we learn how it operates. This chapter represents only a single perspective on this issue. Consequently, we must continue the process of educating one another. That is the goal of this Handbook.
REFERENCES


