11

Cognition

Arthur C. Graesser, Morton A. Gernsbacher
and Susan R. Goldman

When people comprehend discourse, the speech or printed messages are not merely copied into their minds. Instead, the human mind actively constructs various types of *cognitive representations* (that is, codes, features, meanings, structured sets of elements) that interpret the linguistic input. These cognitive representations may incorporate words, syntax, sentetial semantics, speech acts, dialogue patterns, rhetorical structures, pragmatics real and imaginary worlds, and many other levels discussed in this volume. Each type of cognitive representation is functionally important during the processes of comprehending and producing text and talk.

During the last 25 years, cognitive psychologists have explored how the human mind represents the information in various types of cognitive representations. Cognitive psychologists have discovered that some of these cognitive representations are not equivalent to the symbolic representations that have been proposed by many formal linguists, logicians, and computer scientists. For example, suppose that a husband and wife are in the middle of a heated argument and the wife dramatically exclaims, 'If you don't leave, my clothes are going to Boston!' A traditional logician would construct a 'truth table' that specifies the truth values of all combinations of the husband's leaving versus not leaving, and of the clothes going versus not going to Boston. In an effort to 'comprehend' this speech act, a computer program would expend some processing time sorting out exactly which clothes might end up going to Boston. Both the logician and the computer program would miss the important inference that the wife would also be going to Boston (if the husband doesn't leave). In contrast, the meaning representations constructed by humans would not include the entire truth table and the precise set of clothes, but they probably would include the inference that the wife would be leaving. The meaning representations in the human mind are quite elaborate because they are anchored in a rich body of experiences and background world knowledge (which varies from person to person). At the same time, the meaning representations frequently are fragmentary (rather than complete), vague (rather than precise), redundant, open-ended, and sketchy. And yet, with all this apparent slop in the system, writers/speakers manage to

construct messages that frequently can be recovered by readers/listeners with impressive accuracy.

Cognitive psychologists also investigate the mental processes that construct the cognitive representations. Some of these cognitive processes include *accessing* words in the mental lexicon, *activating* concepts in long-term memory, *searching* for information, *comparing* structures that are available in working memory, and *building* structures by adding, deleting, rearranging, or connecting information. Some cognitive processes are executed automatically and unconsciously, at lightning speed (measured in milliseconds). The execution of other cognitive processes is deliberate, conscious, and slow (measured in seconds). Of course, there is a continuum between these two extremes.

How do c gnitive psychologists know whether humans actually construct these cognitive representations and perform these cognitive processes? Psychologists test hypotheses about cognition by conducting experiments and collecting data from humans. For example, there are a number of ways to test whether a reader constructs a particular representation. A group of readers might recall a text after they finish comprehending it. The content that is recalled should to some extent resemble the conitive representations. If a theory predicts that text statement A is more central to the cognitive representation than statement B, then the likelihood that readers later recall A should be higher than that of B. As an alternative method, a series of test statements would be presented after comprehension and readers would decide whether each test statement was explicitly stated in the text. Readers should answer 'yes' when making these decisions to the extent that the test statements match cognitive representations. 'Yes' decisions should frequently occur when a test statement matches an inference that was never stated explicitly. Other tasks that unveil cognitive representations include summary protocols, true/false judgments about test statements, importance ratings for test statements, ratings on the extent to which two statements are conceptually related, and question answering.

It is possible to trace the dynamic process of constructing the cognitive representations 'on-line' during comprehension. One way we do this is by interrupting the reader and collecting data. For example, comprehenders may be asked to 'think aloud' while reading a text. The ideas that come to mind while reading include much of the content that enters the reader's consciousness at particular points in the text. The unconscious mind can also be tapped with experimental tasks. For example, readers might be periodically interrupted during comprehension and be presented test words to name as quickly as possible. The time that it takes to name a test word should be quick if the word closely matches a representation that is active in the reader's unconscious mind. There are experimental tasks that are less disruptive of comprehension than the think aloud task and the word naming task. In eye tracking studies, the researcher records the eye movements and the amount of time that the reader gazes on particular words. Alternatively, self-paced reading times are collected by having readers comprehend text at their own pace; readers press a response button that advances successive text segments one at a time, for example, word by word or sentence by sentence. Reading times for the various text segments are the data to be explained in these self-paced reading time tasks. Cognitive psychologists have devised dozens of experimental tasks that test for the existence of cognitive representations and 'on-line' cognitive processes.

The ultimate goal of the cognitive enterprise is to develop theories that specify how the cognitive representations are constructed and used. These theories are typically complex, given that discourse involves multiple levels and processing components. Moreover, psychological theories of discourse comprehension and production must be grounded in general theories of cognition. A general theory of cognition would explain memory, learning, decision making, problem-solving, and other cognitive faculties in addition to language and discourse. When theories of discourse processing become complex and sophisticated, cognitive psychologists simulate the mechanisms by developing computer models. A good computer model generates output that closely matches the data collected in psychology experiments.

Background and Development of the Cognitive Approach

Early cognitive theories of discourse were inspired by theories of discourse in other fields, such as text linguistics (van Dijk, 1972; Halliday and Hasan, 1976), artificial intelligence (Schank and Abelson, 1977), and pragmatics (Grice, 1975; Searle, 1969). Cognitive researchers explored whether the representations and claims about discourse in these sister fields provided psychologically plausible accounts of representations and processes in humans. Thus, cognitive researchers appropriately sought the wisdom and insights of other fields. As one might expect, some contributions from these sister fields proved to be valid when tested in psychology experiments, whereas other contributions ended up being blind alleys.

Propositional representations attracted the attention of researchers in early psychological theories of discourse (Clark and Clark, 1977; Kintsch, 1974). A *proposition* is a theoretical unit that contains a *predicate* (for example, main verb, adjective, connective) and one or more *arguments* (for example, nouns, embedded propositions), with each argument having a functional role (for example, agent, patient, object, location). A proposition refers to a strefers to a state, an event, or an action and frequently has a truth value with respect to a real or an imaginary world.

In order to illustrate a propositional representation, consider the excerpt in Table 11.1 from the novel *Einstein's Dreams* by Alan Lightman (1993). A propositional segmentation for the first sentence is presented below the excerpt in Table 11.1. The single sentence contains seven propositions. The predicates in these propositions include verbs (lift, place, pinken), adjectives (brown, mushy), and connectives (and, [in order] to). The arguments

Table 11.1 *Excerpt and propositional representation*

Excerpt from Einstein's Dreams (Lightman, 1993: 102)
A mushy, brown peach is lifted from the garbage and placed on the table to pinken. It pinks, it turns hard, it is carried in a shopping sack to the grocer's, put on a shelf, removed and crated, returned to the tree with pink blossoms. In this world, time flows backward.

Propositional representation for the first sentence
Predicates are placed to the left of parentheses; arguments are placed within parentheses. Arguments have functional roles, such as agent, object, and location. PROP stands for proposition.

PROP 1: lift (AGENT = X, OBJECT = peach, SOURCE = from garbage)
PROP 2: brown (OBJECT = peach)
PROP 3: mushy (OBJECT = peach)
PROP 4: place (AGENT = X, OBJECT = peach, LOCATION = on table)
PROP 5: pinken (OBJECT = peach)
PROP 6: [in order] to (PROP 4, PROP 5)
PROP 7: and (PROP 1, PROP 4)

include objects (peach, garbage, table), an unidentified person (X), and embedded propositions (for example, propositions 4 and 5 are embedded in proposition 6). The arguments occupy various functional roles: agent, object, source, location.

Propositions were regarded as the primary functional units for segmenting text. It is important to note that some features of discourse are not explicitly captured in the propositional representations, such as tense, aspect, voice, and the determinacy of nouns. For example, the fact that the example sentence is in the passive voice rather than the active voice is not captured. These auxiliary linguistic features were regarded as comparatively unimportant in the meaning representation of text (Kintsch, 1974).

Cognitive psychologists conducted experiments to test the psychological plausibility of propositional representations. Kintsch (1974) reported that reading times increase as a function of the number of propositions in the text. This trend persists even when there is control over the number of words in the text and many other factors that potentially increase reading times (Haberlandt and Graesser, 1985). Kintsch (1974) reported that recall is better for those propositions that are structurally superordinate (that is, high in a hierarchical tree structure) than those that are comparatively subordinate. For example, propositions 2 and 3 in Table 11.1 are subordinate to proposition 1 because they modify the argument *peach* in proposition 1. Proposition 1 would therefore be recalled more often than propositions 2 and 3; readers would sometimes forget that the peach was brown and mushy.

One challenge for those who advocated propositional theories of discourse was to specify how the propositions are interrelated in a coherent fashion. Obviously, there is an important difference between texts with propositions that fit together conceptually (that is, high coherence) and texts with propositions that are unrelated (that is, no coherence). Kintsch and van Dijk developed psychological models that identified different types of coherenc . Their models specified how coherent text structures are constructed in a working memory with limited capacity (Kintsch and van Dijk, 1978; van Dijk and Kintsch, 1983). It was widely acknowledged that the working memory of humans is limited in capacity. Only a handful of propositions and arguments are available in working memory at any one point in time during comprehension

One level of coherence, called the text *microstructure* (Kintsch and van Dijk, 1978), connects explicit text propositions by argument overlap and other conceptual criteria. Two propositions are linked by *argument overlap* if they share a common argument. For example, propositions 1 and 4 are connected because they share two arguments (X and peach). Sometimes bridging inferences are needed to match arguments of propositions. For example, the pronoun *it* in the second sentence of Table 11.1 refers to *peach* in proposition 1, so there would need to be a bridging inference to capture the overlap: refers-to (it, peach). The process of constructing this bridging inference takes extra processing time to complete (Haviland and Clark, 1974). Although argument overlap was found to be an important criterion for establishing local coherence in many psychological experiments, argument overlap is not the only criterion for connecting propositions at the level of text microstructure (van Dijk and Kintsch, 1983). Local connections are also established by virtue of the situation described by a text, that is the mental microworld. For example, propositions may be connected by relations that convey temporality (proposition A occurred before proposition B), causality (A caused or enabled B), and other dimensions of the microworld. Local connections are established by various types of functional relations between propositions, such as comparison, contrast, generalization, example, and explanation (see Meyer, 1975). These relations, together with argument overlap, provide local text coherence at the microstructure level.

A second level of coherence consists of text *macrostructure.* Text macrostructure interrelates larger segments of text by virtue of world knowledge and genre schemata. For example, a schema for FRUIT DISTRIBUTION would connect many of the events in the example excerpt. One interesting property of this text is that the events are presented in an order that is opposite to the order of events in FRUIT DISTRIBUTION. The reversed ordering is explained by the major point that time flows backward in the imaginary world. The rest of the story shows how this backward flow of time provides illuminating insights about life and reality.

The global schemataat the macrostructure level were vigorously investigated by cognitive psychologists because they were an important key to solving the problem of text coherence. Indeed, text microstructure was hardly sufficient for establishing coherence between propositions. Some global schemata consisted of natural packages of generic world knowledge, such as person stereotypes, object concepts, and scripts (Schank and

Abelson, 1977). A RESTAURANT script, for example, contains knowledge about typical actors (for example, customer, waitress, cook), props (table, menu, food), goals (customer get food, waitress get money), actions (customer sits down, customer orders food, waitress brings food to customer, customer eats, etc.). The generic RESTAURANT script would have tentacles to many propositions in a text about a restaurant and would thereby provide global coherence. The script would supply the world knowledge that is needed for the reader to generate expectations, interpret incoming propositions, and generate inferences. Typical script content is filled in inferentially, which makes it difficult for the comprehender to determine whether a typical script proposition was explicitly stated or merely inferred by default (Bower et al., 1979; Graesser et al., 1979). According to the recognition memory experiments reported by Graesser et al. (1979), adults are entirely unable to discriminate whether a very typical script action (such as eating food) is explicitly stated or merely inferred by virtue of the RESTAURANT script.

Another class of global schemata is associated with particular text genres. Texts can be broadly classified into four different genres: descriptive, narrative, expository, and persuasive. There are various subclasses within these broad categories, and some texts are hybrids of multiple genres. Cognitive psychologists initially spent most of their efforts analysing expository texts (Kintsch, 1974; Meyer, 1975) and simple stories (Mandler and Johnson, 1977; Stein and Glenn, 1979). For example, the schema for a simple folktale has a set of components (for example, characters, setting, plot, episodes, resolution) and permissible orderings of these components (for example, the setting comes before the plot).

A different foundation for analysing text coherence addresses a *given–new* distinction (Haviland and Clark, 1974). An incoming sentence in a text contains both *given* information (that is, a proposition or argument already mentioned in the text) and *new* information. When the incoming sentence is interpreted, the comprehender first searches the previous passage context for information that matches the given information. If a match is found, the new information is appended structurally to the old proposition or argument. It takes a longer time to *reinstate* a proposition that was read several sentences earlier than to refer to a proposition that is resident in working memory. If no match is found, then a new structure needs to be built. Later in this chapter we will discuss a more recent model, called the structure building framework (Gernsbacher, 1990), that expands the given–new distinction.

Early cognitive models of discourse had a heavy emphasis on properties of the explicit text. That is, researchers proposed a quasiformal system for segmenting and organizing text, and then investigated whether these representations explained data in psychology experiments. By the early 1980s, cognitive researchers had identified some limitations with this preoccupation with the explicit text. They seriously acknowledged the importance of the reader and the constraints of general cognition. Comprehension came to be viewed as an active, flexible, strategic process rather than a passive, inflexible translation of explicit code (Graesser, 1981; van Dijk and Kintsch, 1983). It was important to consider the goals and background knowledge of the reader. Why was the reader comprehending the text? Was it read for entertainment, for a later memory test, or for proofreading? What did the reader know about the topic being discussed? Was the reader an expert or a novice about the topic? These reader characteristics profoundly influenced the cognitive representations constructed by comprehenders (Spilich et al., 1979). Although cognitive psychologists were always aware that meaning did not reside exclusively in the text *per se*, it was time to consider the reader characteristics more seriously. There was an increasing concern for the representation of the world knowledge that was activated by the text and for knowledge-based inferences (Graesser and Bower, 1990).

So far, our discussion of early psychological research on discourse has focused on reading (rather than talking) and has ignored the fact that most discourse is designed to communicate ideas in a social context. In fact, however, cognitive psychologists were quite aware of the communicative, social, and pragmatic dimensions of discourse (Bates, 1976; Clark and Clark, 1977). It was widely acknowledged that discourse comprehension and production are embedded in a communication system with three components: the writer/speaker, the reader/listener, and the text/talk. The speakers and listeners are visible, specific, and co-present in conversations. The speech acts in conversations occur in a specific context, situation, location, and time span. The speech participants have some sense of what knowledge they share (called common ground or mutual knowledge) and what goals they are attempting to achieve in the conceptually rich, situated context (Clark and Schaefer, 1989). In contrast, the writers, readers, and written texts are normally *decontextualized*. That is, the writer is invisible to the readers and the readers are invisible to the writer. Written text is produced at a different context, situation, location, and time than the comprehension of the text. The writer and reader are not always privy to what each other knows and what their separate goals are. However, in spite of these differences between text and talk, discourse is still embedded in a communication system with the three components.

Speech acts were the basic unit of linguistic analysis for those psychologists who concentrated on conversation analysis. According to the speech act theories, the stream of conversation is segmented into speech acts (D'Andrade and Wish, 1985; Searle, 1969). The representation of each speech act is a complex description that varies somewhat among speech act theorists. A speech act description might specify the speaker, the addressee, the literal propositional content, the speech act category (for example, assertion, question, promise, threat, request), and the intended meaning. For example, consider once again the wife expressing the following speech act to her husband: 'If you don't leave, my clothes are going to Boston!' The speaker is the wife, the addressee is the husband, and the speech act

category is that of a threat. The literal propositional content is the conditional expression (if–then), with two embedded propositions (one about the husband not leaving and the other about the clothes going to Boston). The intended meaning stipulates that the wife plans on leaving if the husband doesn't leave.

It is important to note that there does not need to be a high semantic similarity between the literal meaning of a speech act (that is, its propositional content) and the intended meaning of the speech act (Searle, 1969). Suppose that you are at a dinner table and a person asks, 'Could you pass the salt?' This speech act is intended as an *indirect request* for you to pass the salt, even though it is literally expressed as a question about your salt passing abilities. Suppose that there is an angry storm outside and a friend of yours comments, 'Lovely weather outside.' This is an *ironic* utterance because the literal meaning is opposite of the intended meaning.

Clark and Lucy (1975) once proposed a two-stage model to account for the time-course of comprehending speech acts that involved discrepancies between the intended meaning and the literal meaning. The speaker first constructs the literal meaning and then, after detecting problems with the literal meaning, constructs the intended meaning. Therefore, extra processing time is needed to construct the intended meaning. Subsequent research challenged the two-stage model. Extra processing time was not necessarily needed to recover the intended nonliteral meaning. Instead, intended meanings can be constructed quickly (Glucksberg et al., 1982). One of the lively contemporary debates addresses the process of constructing intended meanings of speech acts on the basis of context plus the explicit text (Gibbs, 1994).

Current Directions, Theories, and Phenomena Investigated

Cognitive studies of discourse have flourished during the last 25 years. Researchers have published dozens of books and hundreds of articles in approximately a dozen different journals. Space limitations in this chapter do not permit a comprehensive treatment of all of the exciting research trends, phenomena, and theories. We instead focus on those topics that have received substantial attention in cognitive psychology and that also intersect our own programs of research. Consequently, this chapter covers text comprehension to a greater extent than text production and conversational discourse.

Cognitive Models of Discourse

Cognitive psychologists have been quite persistent in building sophisticated models of cognitive mechanisms. These models specify the representations, processing components, and interactive mechanisms in enough detail that patterns of empirical data can be simulated. Computational models simulate cognitive mechanisms on a computer. Mathematical models quantify precise patterns of processing times, memory scores, ratings, and other psychological data. These computational and mathematical models exist in the arena of discourse (Britton and Graesser, 1995; Just and Carpenter, 1992; Kintsch, 1988; Weaver et al., 1995), just as they do in other areas of cognitive psychology. It should be noted that modeling efforts are useful even when the simulated output fails to match human output. An understanding of why such discrepancies occur unveils new insights about the limitations of existing models and provides some direction for further research.

Psychological models of discourse have been greatly influenced by two major cognitive theories: symbolic theories and connectionist theories. In symbolic theories (Anderson, 1983), there is a working memory (as discussed earlier) and a vast storehouse of concepts, propositions, schemata, and production rules. A *production rule* has an 'IF [conditions] THEN [action]' format. When the conditions are met, the production is 'fired' (that is, activated) and the action (or action sequence) is performed. For example, the following simple production rule occurs frequently in most households:

IF [a telephone rings and a person is near the telephone]
THEN [the person picks up the telephone and says 'hello']

Production rules may involve cognitive actions rather than physical actions:

IF [the letter sequence h-e-r-o is percieved]
THEN [activate the concept of HERO in working memory]

A production system has thousands of these production rules. The production rules are continually being evaluated during each cycle of comprehension. According to some models, there are dozens or hundreds of these cycles of comprehension during a mere second. As new input enters working memory, all production rules are evaluated in parallel, but only a few of the production rules are fired, namely those that have their conditions satisfied. As new production rules are fired, and the information in working memory changes, verbal or physical actions are produced as output. The information in working memory dynamically changes over time, from cycle to cycle, as dictated by perceptual input and the knowledge base in long-term memory. The system learns from these dynamic changes in working memory. The process of learning creates new facts and production rules in long-term memory.

In connectionist theories (McClelland and Rumelhart, 1986), representations and processes are distributed among a large set of simple units. The units are often called *neural units* because there is a metaphor with neurons in the brain. Intelligent activity is believed to emerge from a large, interconnected mass of simple neural units. Each word, proposition, concept, schema, or rule has a corresponding ensemble of neural units. The activation level of each unit fluctuates dynamically over time, as comprehension proceeds. The units are connected by *weights*, thereby forming a *neural network*. The weight that connects one unit to another unit may be either excitatory (positive weight), inhibitory (negative weight), or zero. In a fully connected network, each unit is connected to every other unit (including itself). Therefore, if there are N units, there would be $N \times N$ weights in the *weight space*. The knowledge in long-term memory consists of the set of units and the weights in the weight space. When learning occurs, there is a change in one or more of the weights in the weight space.

So what happens during a particular comprehension cycle? A set of units is initially activated, namely those that capture the context and the perceived input. These units then excite or inhibit their neighboring units, according to the weights in the weight space; the neighbors then activate or inhibit their neighbors, and so on. Eventually, stability is achieved in the network when there are minimal changes in the activation values of the units: the network settles into a stable *pattern of activation*. The meaning representation at a particular point of comprehension consists of the pattern of activation values for all units. In this sense, meaning is said to be *distributed* throughout the network. In contrast, in a symbolic system, meaning is localized to one or a few symbolic expressions.

Although there have ben a few bona fide connectionist models of text and discourse (for example, St John, 1992), most models are hybrids of the symbolic and connectionist theories (Britton and Graesser, 1995; Golden and Rumelhart, 1993; Goldman and Varma, 1995; Just and Carpenter, 1992; Kintsch, 1988). At this point, we will briefly describe the two most influential models of comprehension in cognitive psychology: the construction-integration model (Kintsch, 1988) and the collaborative activation-based production system model (Just and Carpenter, 1992).

Construction-Integration (CI) Model Kintsch's (1988) CI model distinguishes three levels of representation: the surface form, the propositional textbase, and the referential situation model. The surface form preserves the exact words and syntax of sentences, whereas the textbase is similar to the propositional microstructure that was described earlier (see Table 11.1). The situation model integrates the text information with the reader's world knowledge and refers to the unique world that is conveyed in the text.

The CI is a hybrid model that combines symbolic expressions and connectionistic weights. The symbolic expressions include the content words (that is, nouns, main verbs, adjectives), the explicit text propositions, and world knowledge relevant to the text (which also comprises word and proposition expressions). For example, in the case of the text in Table 11.1, the first sentence would include 10 word units (lift, brown, mushy, place, pinken, [in order] to, and, peach, garbage, table), 7 proposition units, and 2 or more units referring to relevant world knowledge (for example, FRUIT DISTRIBUTION, grocer, other information that will not be specified here). These 19 units (10 + 7 + 2) are connected by a set of 19×19 weights, in the spirit of connectionist models. The weights are specified theoretically according to the constraints of the surface form, the textbase, and world knowledge. Consider the theoretical weight space corresponding to the textbase. Proposition 1 would have a positive weight connecting to proposition 2 by virtue of argument overlap; however, proposition 6 would not be directly connected to proposition 1 because there is no direct argument overlap in that case. Consider the weight space that involves the situation model. There would be a positive weight between grocer and proposition 1, signifying that grocer is the likely agent that lifts peaches from the garbage; there would not be a positive weight between grocer and proposition 5. Therefore, the CI model has a separate weight space for t he surface form, the textbase, and the situation model. Each weight space has the same 19 nodes (and of course others that we will not bother mentioning).

The CI model s mulates the dynamic fluctuation of activation values for the units in the network. These values change as comprehension proceeds, word by word, proposition by proposition, and sentence by sentence. At each cycle of comprehension, new words activate some of the units, activation spreads through the network, and the pattern of activation values for units eventually stabilizes. Then a new cycle of comprehension occurs and the process starts all over again. As a consequence, one can observe the activiation value of each unit as a function of the sequence of the comprehension cycles.

The construction phase of the CI model consists of the creation of units corresponding to the explicit text and the associated world knowledge. These units, plus units from the prior discourse, are activated to varying degrees. The *integration* phase is the process of settling on a stable pattern of activation values. On the average, the units that have positive connection weights to many other units will settle on high activation values; units that are detached from other units will have low activation values. Therefore, coherence among the units is achieved in a systematic manner, but the connections have strength values rather than being discrete (that is, all-or-none).

Working memory plays an important role in the CI model. The CI model assumes that there are limitations on the amount of information that can be active in working memory at any point in t me. Working memory holds the current sentence being processed and a set of propositions that is carried over from the previous comprehension cycle. The number of propositions carried over is designated as parameter *s* (designating size). The selected propositions are those that have the highest activation value. The value for *s* has been 2 in most of Kintsch's simulation efforts, but in principle this value could vary. The important assumption is that there is a fixed-capacity buffer. Those propositions that are not carried over in workig memory still remain in long-term memory. However, these stored propositions can be reinstated in working memory if they are activated once again in a subsequent comprehension cycle.

The CI model predicts patterns of data in psychology experiments (Goldman and Varma, 1995; Kintsch, 1988; Haenggi et al., 1995). For example, when readers are asked to recall the text after comprehension, the likelihood of recalling the various propositions differs substantially. These recall likelihoods are positively correlated with their average activation values across the comprehension cycles. In some experiments, readers are stopped at the end of a sentence and presented with a letter string that either does or does not form a word (for example, grocer versus croger); readers decide as quickly as possible whether the test string does or does not form a word by pressing one of two buttons. The speed of these *lexical decisions* is correlated with a word's activation value, as computed by the CI model. For example, the lexical decision speed for 'grocer' would be facilitated even though it was never explicitly mentioned; it would have been activated in the situation model

Collaborative Activation-Based Production System (CAPS) Model The CAPS model is also a hybrid between the symbolic and connectionist theories (Just and Carpenter, 1992). There are symbolic expressions, such as words, phrases, propositions, schemata, and production rules. The information in working memory dynamically changes as production rules are fired in response to input. Unlike many other production systems, however, the conditions of production rules can exist at varying degrees of ativation rather than being present versus absent (that is, all-or-none). The condition for a production rule is satisfied if the total activation value meets or exceeds some threshold. Consider the earlier production rule that activated the concept of HERO in working memory when the letters h, e, r, and o were registered as four elements in the condition. Suppose that the overall threshold for activating the rule is 100 units of activation in the condition. The production rule would fire if the activation values for h, e, r, and o are 40, 40, 0, and 40, respectively, because the total activation is 120, which exceeds the threshold. Thus, it would not be essential to detect all four letters in order to fire the production rule. Like all production rules, when the production rule is fired it performs the specified physical or cognitive processes.

The CAPS model captures the fact that working memory is limited in capacity and this limitation influences comprehension (Daneman and Carpenter, 1980; Whitney et al., 1991). CAPS assumes that there is a limit on the total amount of activation available for working memory elements, called the *cap*. When fired production rules request more activation than is available, the cap has been reached. Processing at the cap results in an overall system slowdown and a graceful loss of those working memory elements that are not participating in the processing. When an element falls below a minimum level of activation, it is no longer functional in working memory and cannot participate in processing.

Just and Carpenter have used the CAPS model to simulate reading times for individual words as readers comprehend sentences in text. The word reading times have been measured by collecting eye tracking data or by collecting self-paced word reading times. Longer reading times are predicted at points in the sentence when the cap is reached. Longer reading times are also predicted when the interpretation of an incoming word requires several microcycles of processing. Just and Carpenter have reported that the reading times for individual words are sometimes sensitive to the working memory spans of individual readers and that CAPS can account for the different patterns of reading times in high- versus low-span readers (Just and Carpenter, 1992; Millis and Just, 1994).

CI and CAPS Together Goldman and Varma (1995) developed a model that combines features of the CI model and the CAPS model. The fixed-buffer working memory of the CI model was replaced with a CAPS method of allocating ativation in working memory. As a consequence, instead of carrying over only s propositions to the next comprehension cycle (as in the CI model, where *s* is normally 2), there is a more complex and judicious selection of proposition units to carry over. The improved model by Goldman and Varma provided a longer passage history, more interconnections among propositions, and an enhanced formation of global macrostructures than did the CI model. Goldman and Varma's augmentation of the CI model with the CAPS model corrected one of the disappointing features of the CI model: local microstructure features of the text tended to dominate processing so the simulated reader frequently ended up losing the big picture. One other advantage of Goldman and Varma's hybrid model is that it integrated the goals and strategies of the reader into the comprehension mechanism. It was beyond the scope of the CI model to handle the systematic repercussions of particular reader goals and strategies.

Interactive Processing of Multiple Levels of Discourse and Knowledge

Everyone agrees that discourse comprehension involves multiple, highly interactive components. However, there have been some heated debates about the nature and timing of these interactive processes. For example, suppose that the reader encounters the word *he* in the middle of a novel. The process of resolving the referent of the pronoun would be influenced by sentence syntax, local semantic constraints, and the protagonists that exist in the discourse focus. Would syntax, semantics, or discourse focus have the most robust impact on fetching the correct referent for *he*? Which levels of analysis would be executed most quickly?

According to modularity theory (Fodor, 1983), there is an autonomous module for processing syntax and this module is more quickly executed than local semantics and discourse components. Discourse and semantics may subsequently override the syntax module, but it is syntax that reigns supreme early in the processing stream. For example, suppose that a reader encounters the sentence 'The thief stopped the girl with the dress.' According to a highly regarded hypothesis about syntax, called the *minimal attachment hypothesis* (Frazier and Fodor, 1978), there is a preference to construct syntactic structures that add a small number of new nodes to the syntactic representation. The theory predicts that the syntactic component would have an initial bias to interpret 'with the dress' as an instrumental prepositional phrase that elaborates the activity of 'stopping'. The local semantic context would later override this interpretation by assigning 'with the dress' the status of a relative clause that modifies 'girl'. So syntax is always executed first, even though semantics and discourse later prevail.

According to *interactive theories*, however, modules are highly interactive (rather than autonomous) and there is no intrinsic ordering of syntax before semantics and discourse (Just and Carpenter, 1992). Sometimes the constraints of discourse reign supreme and have a swift impact on the processing of a word, compared to the impact of local semantics and syntax (Hess et al., 1995). At other times, the local semantic context reigns supreme because it is more constraining. In addition to the modularity and interactive positions, there are a host of other models that specify interactions among lexical, syntactic, semantic, and discourse components (Perfetti, 1990).

Nevertheless, cognitive researchers do agree that comprehension is not a completely bottom-up process. It is not the case that syntax is initiated and entirely completed before semantics begins, or that semantics is completed before discourse processes are initiated. Instead, partial analyses evolve at all levels until the final representation is achieved. A computational model is *not* psychologically plausible if it requires a complete and accurate analysis of one component N before proceeding to another component M.

On-Line Construction of Coherent Representations

Cognitive researchers have investigated the process of constructing coherent representations at different levels of discourse during on-line comprehension (Gernsbacher, 1990; Lorch and O'Brien, 1995; Zwaan et al., 1995). Comprehension time for an incoming proposition is comparatively fast if it matches a proposition in working memory (an explicit proposition or an inference) and a bit longer if it appends new information to a proposition in working memory. Comprehension time increases if the reader needs to reinstate information mentioned earlier in the text and that no longer resides in working memory. Comprehension time increases to the extent that inferences must be made to connect the incoming sentence to prior text. A proposition takes a long time to comprehend if it is not related to any information in working memory and the previous context; in these instances, the reader builds a new structure and sometimes regards the information as irrelevant.

Strata of Meaning While reading a story, coherence is potentially monitored on several strata of meaning. These include: (1) the overlapping arguments in propositions, (2) the spatial locations of entities, (3) the causal flow of events, (4) the goals and plans of protagonists, (5) the temporal chronology of episodes (6) the main point or theme, and (7) the purpose of the author in expressing a particular proposition. Cognitive psychologists have investigated the process of constructing representations at each of these strata, but particularly argument overlap (Kintsch and van Dijk, 1978), spatiality (Haenggi et al., 1995; Morrow et al., 1987; Rinck and Bower, 1995), causality (Fletcher and Bloom, 1988; Myers et al., 1987; van den Broek and Lorch, 1993), the goals and plans of characters (Dopkins et al., 1993; Long et al., 1992; Trabasso and Suh, 1993), and temporality (Zwaan et al., 1995).

Under what conditions do comprehenders monitor these various strata? Cognitive psychologists are divided on the answer to this question. According to McKoon and Ratcliff (1992), argument overlap among propositions is a critica stratum to monitor; coherence and elaborative inferences are monitored at the other strata only if there is break in argument overlap or if the reader has comprehension goals that are tuned to a particular stratum. According to other researchers, however, lack of argument overlap is neither necessary nor sufficient for relations to be constructed at the other strata (Albrecht and O'Brien, 1993; van den Broek and Lorch, 1993; Zwaan et al., 1995). Instead, several strata are simultaneously monitored.

Structure Building Framework Gernsbacher (1990) proposed this model to account for the process of building coherent cognitive representations on-line. The process of building structures involves a number of subprocesses. First, comprehenders *lay foundations* for the mental structures. Next, comprehenders develop the structures by *mapping on information* when that information is related to the previous information. When the incoming information is less coherent or related, comprehenders employ a different process: they shift to *initiate a new substructure*. Therefore, most representations have several branching substructures.

The building blocks of the mental structures are called *memory nodes.* These nodes are activated by incoming stimuli and they transmit processing signals to other nodes. The processing signals either enhance (boost) or suppress (dampen) the activation levels of other nodes, much in the spirit of connectionist models. Memory nodes are enhanced when the information represented is necessary for further structure building. Nodes are suppressed when they are no longer necessary for building the multi-leafed structures.

Gernsbacher (1990) and her colleagues have extensively investigated the three subprocesses of structure building: (1) laying a foundation, (2) mapping relevant information onto the foundation, and (3) shifting to initiate a new substructure. The first two processes explain a persistent empirical phenomenon called the 'advantage of first mention'. That is, participants mentioned first in a sentence are more memorable than participants mentioned later. For example, after comprehending the sentence 'Tina beat Lisa in the state tennis match', Tina would be more memorable

and quicker to access than would Lisa. The first participant is normally the discourse topic and lays the foundation of the mental structure. In a series of experiments, Gernsbacher ruled out potential extraneous explanations of this empirical finding, such as the fact that Tina is an agent and the syntactic subject. The empirical finding occurs even in those languages, such as Spanish, where the order of words is less constrained than in English.

The occurrence of processes 1 and 3 explains a second persistent empirical phenomenon, called the 'advantage of clause recency'. That is, information in the most recent clause in a sentence is more memorable and accessible than information from an earlier clause in the sentence. For example, the word 'oil' is more accessible immediately after comprehending sentence 1 than sentence 2:

1 Now that artists are working fewer hours, *oil* prints are rare.
2 Now that artists are working in *oil*, prints are rare.

Comprehenders represent each clause of these two-clause sentences in its own mental substructure. While building a clause-level substructure, comprehenders have the greatest access to information in that substructure. However, after a comprehender has finished building a representation for the most recent clause, information from the first clause becomes more accessible becasue it is the foundation for the entire sentence. Consequently, the advantage of first mention is a long-lived phenomenon whereas the advantage of clause recency is short-lived.

Gernsbacher identified some of the discourse cues that encourage the second process of mapping. The explicitness of a referring expression is an important cue for signaling coherence, as has been observed by linguists (Givón, 1993; Halliday and Hasan, 1976). *A big bad wolf* is an indefinite noun-phrase so it would signal a new structure. In contrast, *the wolf* would probably signal a mapping of information onto an existing structure and the pronoun *it* would signal mapping rather than shifting. Causal coherence involves mapping (rather than shifting) and is frequently signaled by connectives, such as *because*, *so*, and *in order to.* Temporal mapping is frequently signaled by the tense and aspects of verbs. It should be noted that surface codes and function words in a clause become less accessible whenever there is a shift to a new substructure during process 3.

Gernsbacher and her colleagues have examined the processes of enhancing and surpressing nodes during comprehension. When comprehenders encounter homographs (such as 'spade'), multiple meanings are immediately activated (for example, garden tool versus card suit) even though one meaning is appropriate for the context (such as 'He dug in the garden with a spade'). However, within a half a second, only the contextually appropriate meaning is available. What happens to the contextually inappropriate meanings? According to the structure building framework, they do not receive lower activation by mere decay or by competitive inhibition among alternative meanings. Instead, they are actively suppressed by the signals transmitted by memory nodes that represent the syntactic, semantic, and pragmatic context. Less skilled readers have problems suppressing the inappropriate meanings of words whereas skilled readers have efficient suppression mechanisms (Gernsbacher, 1993). Comprehension skill is predicted by the quality of the suppression mechanism but not by the enhancement mechanism.

Constructing Inferences and Situation Models

As discussed earlier, van Dijk and Kintsch (1983) contrasted three levels of representation: the surface form, the propositional textbase, and the referential situation model. The situation model refers to the people, setting, states, events, and actions of the mental microworld that the text describes. For example, in the text in Table 11.1, the reader would imagine the process of the peaches ripening and being distributed, but in reversed order, much like a videotape going backward. Situation models have been difficult to investigate systematically for several reasons. First, world knowledge plays a central role in building situation models, yet world knowledge is imprecise, open-ended, vague, and minimally visible to the investigator. Second, situation models are unique representations that embody the idiosyncratic constraints of the particular text. It is difficult to identify general mechanisms when the representations are so idiosyncratic. Third, there is no sophisticated theory from any field that both specifies how situation models are constructed and also is psychologically plausible. Psychologists have been forced to discover the mechanisms on their own. This has presented a stimulating challenge to many cognitive psychologists. There have been lively debates over what classes of inferences are generated during reading and the construction of situation models (Graesser and Bower, 1990; Graesser et al., 1994; McKoon and Ratcliff, 1992).

Narrative text has received the most attention in studies of inference generation and situation model construction. This is because narrative texts embody episodes that resemble everyday experiences and that activate an extensive mass of world knowledge. In contrast, expository texts typically inform the reader about topics that the reader is unfamiliar with. Therefore, cognitive researchers have examined whether different classes of inferences are generated on-line during the comprehension of narratives. These classes of inferences include character traits, the knowledge and beliefs of characters, the goals and plans that motivate character actions, the manner of executing actions, the spatial setting and layout, the causes of events, emotional reactions of characters, and expectations about future episodes in the plot. To a lesser extent, researchers have explored the inferences associated with the pragmatic interaction between writer and reader, such as inferences about the attitudes of the writer and appropriate emotional reactions of the reader. Although all of these classes of inferences could potentially be generated in an elaborate analysis of a story, only a subset of the inferences is generated on-line during an initial reading of a text. A

Table 11.2 *The czar story and example inferences*

The Czar and his Daughters

Once there was a czar who had three lovely daughters. One day the three daughters went walking in the woods. They were enjoying themselves so much that they forgot the time and stayed too long. A dragon kidnapped the three daughters. As they were being dragged off, they cried for help. Three heroes heard their cries and set off to rescue the daughters. The heroes came and fought the dragon and rescued the maidens. Then the heroes returned the daughters to their palace. When the czar heard of the rescue, he rewarded the heroes.

Inferences when comprehending 'The dragon kidnapped the daughters'

SUPERORDINATE GOAL:	*The dragon wanted to eat the daughters.*
A goal that motivates an agent's intentional action.	
SUBORDINATE GOAL/ACTION:	*The dragon grabbed the daughters.*
A goal, plan, or action that specifies how an action is achieved.	
CAUSAL ANTECEDENT:	*The dragon saw the daughters.*
An event or state on a causal chain that bridges an explicit proposition to the previous passage context.	
CAUSAL CONSEQUENCE:	*Someone rescued the daughters.*
A physical event or action on a forecasted causal chain that unfolds from an explicit proposition. Emotional reactions in characters are not included.	
EMOTIONAL REACTION:	*The daughters were frightened.*
An emotion experienced by a character in response to an explicit event, action, or state.	
STATE:	*The dragon has scales.*
An ongoing state, from the time frame of the story plot, that is not causally linked to episodes in the plot. These include character traits, properties of objects, and spatial relationships among entities.	

good theory should be able to discriminate inferences generated on-line versus off-line.

Table 11.2 presents an example story about a dragon kidnapping three daughters and being saved by heroes. There are definitions and examples of six classes of inferences: *superordinate goals* that motivate characters' actions, *subordinate goals* that specify how actions are achieved, *causal antecedents* of events, *causal consequences*, *emotional reactions* of characters, and ongoing *states*. These example inferences are *extratextual* inferences rather than *text-connecting* inferences. Text-connecting inferences specify that two or more explicit propositions are connected conceptually. Extratextual inferences embellish the situation model by copying or deriving information from world knowledge. An investigation of the inferences illustrated in Table 11.2 has provided informative tests among different models of inference generation. The models make different predictions about which classes of inferences are generated on-line.

Explicit Textbase Position This position is compatible with early models of text comprehension that focused on the explicit text (Kintsch, 1974; Mandler and Johnson, 1977). According to this position, the explicit textbase reigns supreme in shaping the cognitive representation of discourse, not the situation model. The only inferences that are constructed on-line are the referential inferences that bind explicit arguments and propositions in text (for example, linking pronouns to previous arguments, establishing argument overlap). None of the extratextual inferences in Table 11.2 are constructed on-line according to this first position.

Minimalist Hypothesis McKoon and Ratcliff (1992) proposed this hypothesis to account for those inferences that are automatically (versus strategically) encoded during comprehension. The only inferences that are encoded automatically during reading are those that make text statements locally coherent. Situation-based inferences are encoded only when there is a break in local coherence (specifically, argument overlap) or when the reader has a goal to construct a particular class of inferences (for example, the goal of tracking the spatial locations of characters and objects). Causal antecedent inferences are the only inferences in Table 11.2 that are important for establishing local text coherence; readers need to construct causal antecedents in order to causally bridge an incoming story event with the prior passage context. Therefore, the minimalist hypothesis predicts that causal antecedent inferences should have the highest strength of encoding during comprehension and that the other inferences are sporadically generated on-line.

Current-State Selection Strategy and the Causal Inference Maker Model These models specify the process of constructing causal connections between explicit actions and events in stories (Fletcher and Bloom, 1988; van den Broek and Lorch, 1993). According to these two models, only two classes of extratextual inferences are reliably generated on-line: causal antecedents and subordinate goals. However, it is beyond the scope of this chapter to describe the mechanisms that supply these predictions.

Constructionist Theory Graesser et al. (1994) developed a constructionist theory that has three major assumptions. The *reader goal* assumption states that comprehenders construct inferences that address the comprehenders' goals. This first assumption does not offer any discriminating, invariant predictions about the on-line status of the inferences in Table 11.2, but it does offer context-sensitive predictions that consider the idiosyncratic goals of the reader. The second assumption, the *coherence* assumption, states that comprehenders attempt to construct a meaning representation that is coherent at both local and global levels. Whereas causal antecedents are important for establishing local coherence, superordinate goals and emotional reactions of characters are important for establishing global plot coherence in stories. It should be noted, however, that attempts to construct global coherence constitute an effort, not necessarily an achievement. If the text is choppy, meandering, and pointless, readers will give up trying to construct a globally coherent meaning representation. According to the third, *explanation* assumption, comprehenders attempt to explain *why* actions, events, and states are mentioned in the text. Inferences that

answer such why-questions include causal antecedents and superordinate goals (Graesser, 1981). In summary, the constructionist theory predicts that three classes of inferences are reliably constructed on-line when individuals comprehend stories: causal antecedents, superordinate goals, and emotional reactions of characters. The other three classes of inferences in Table 11.2 are elaborations that are not normally constructed on-line: causal consequences, subordinate goals, and states.

Prediction-Substantiation Model This model asserts that comprehension is expectation-driven in addition to explanation-driven (Bower et al., 1979; Schank and Abelson, 1977). Expectations are formulated when a higher-order knowledge structure is activated (such as a script); the content of the global structure supplies the expectations. The only classes of inferences in Table 11.2 that would *not* be generated on-line are subordinate goals and states. It should be noted that the previous models resisted the possibility that expectation-based consequence inferences are generated on-line. This is because (a) there are too many future plots that could conceivably unfold, (b) it takes a large amount of working memory resources to construct even a single hypo hetical plot, and (c) most expectations end up being disconfirmed in the face of subsequent discourse (Graesser, 1981; Kintsch, 1988).

Promiscuous Inference Generation This extreme position predicts that all six classes of inferences are generated on-line, as long as the reader has the prerequisite world knowledge. Comprehenders build a complete, lifelike situation model by fleshing out all of the details about the characters, props, spatial layout, actions, events, and so on. The meaning representation is akin to a high-resolution mental videotape of the narrative, along with complete information about the mental states of characters. It would be difficult to find a researcher who seriously advocates this position, but it nevertheless is an interesting extreme position to consider.

Tests of the Models of Inference Generation An adequate test of the above models would need to assure that comprehenders have the prerequisite background knowledge to generate the extratextual inferences. It would be pointless to assess whether a class of inferences is generated on-line if few readers had the critical world knowledge. Therefore, some researchers have collected 'think aloud' protocols or question answering protocols from a sample of readers in order to extract potential inferences that may (or may not) be generated on-line during comprehension (Graesser, 1981; Trabasso and Suh, 1993). These verbal protocols are collected as the sample of readers comprehend the text, sentence by sentence. If an inference is manifested in these protocols, there is some assurance that the reader has the prerequisite knowledge to make the inference. However, more rigorous experimental tests are needed to assess whether the inference is truly made on-line during normal comprehension (that is, when verbal protocols are not collected).

One type of experimental test involves interrupting the comprehender at critical points in the text and collecting word naming data or lexical decision data. Suppose that the reader of the story in Table 11.2 is interrupted after reading the explicit statement 'The dragon kidnapped the daughters.' The alternative test words to name would come from the six inference classes: *eat*, *grab*, *see*, *rescue*, *fright*, and *scales*. Of course, the words selected in tests of the various inference classes would be equated on word frequency, syntactic category, number of letters, and a host of other extraneous factors.

Graesser et al. (1994) examined evidence from several dozen experiments on inference generation in order to evaluate the above theoretical positions. They argued that most of the existing evidence supports the constructionist theory rather than the other positions. For example, readers generate superordinate goals on-line but not subordinate goals (Long et al., 1992). They generate causal antecedents but not causal consequences (Magliano et al., 1993; Potts et al., 1988). They generate character emotions (Gernsbacher et al., 1992) but not ongoing states. The constructionist model apparently is compatible with most of the existing empirical evidence, but more research is needed before we can be confident in our claims about inference generation and the construction of situation models.

Comprehending Bona Fide Literature

An adequate model of discourse comprehension would generalize to naturalistic texts rather than being restricted to experimenter-generated 'textoids' (van Oostendorp and Zwaan, 1994). One of the current trends is to explore the process of comprehending actual literary texts, such as short stories and novels (Dixon et al., 1993; Gerrig, 1993; Kreuz and MacNealy, 1995; Miall and Kuiken, 1994; Zwaan et al., 1995).

Literary narrative has a number of properties which are different from the narrative, expository, and pseudotexts that cognitive psychologists have traditionally analysed. Literature is written in part to produce emotional responses in the reader, such as surprise, curiosity, or suspense (Brewer and Ohtsuka, 1988). Literature is written to reveal deep truths about life and reality, even when the plot is entirely fictional. Literature has a high density of nonliteral forms, such as irony, metaphor, understatement, and hyperbole. Literary excerpts are sometimes crafted to support multiple interpretations (that is, intentional ambiguity) rather than to converge on a single intended meaning. Literary texts often violate linguistic and social conventions, thereby encouraging the reader to reflect on language and society. Literary texts have points of view either that are unusual (for example, observing events from the point of view of a dog) or that fluctuate (for example, between a character and an omniscient narrator). Investigations of literary comprehension both open the door to new comprehension phenomena and test the limits of the mainstream cognitive theories of discourse.

The excerpt in Table 11.1 is an example of a literary device. The reader starts out rather confused about the temporal sequence of events in the first two sentences. This confusion exists even though the text exhibits local coherence, namely the argument overlap involving the peach. The third sentence resolves the confusion: time flows backward in that world. From the perspective of the third sentence, all of the events make sense when they are reinterpreted. This device builds confusion, followed by revelation; it creates tension and then abrupt release, which is a common arousal pattern in aesthetic works.

Major Applications

The importance of cognitive research on discourse is not restricted to academic circles. There have been several applications of cognitive models to real world problems. For example, some salient areas of application have been in education, in the designe of computer displas and dialogue facilities, and in the construction of printed texts, documents, questionnaires, and forms.

The cross-fertilization between education and cognitive research on discourse has been one of the productive arenas during the last 10 years. The cognitive enterprise has offered some promising new theories and some rigorous experimental methodologies, whereas the education enterprise has forced the cognitive researchers to investigate naturalistic texts and learning environments. A host of practical questions in education have been explored as a consequence of this cross-fertilization (Barr et al., 1991). How do children read stories and expository texts? How can reading be improved? How do readers of all ages differ in reading abilities? How do readers learn from text, and how can this be improved? How can texts be designed or revised to enhance learning and memory? How is learning from text facilitated by auxiliary organizers, such as outlines, questions, highlighting, pictures, diagrams, and animation?

In the best of worlds, learning materials are designed to maximize reading speed, comprehensibility, information delivery, memory, and enjoyment. Unfortunately, however, there sometimes are tradeoffs in meeting these design goals. For example, a text that is pitched for enjoyment may be lean in information delivery. A text that is difficult to comprehend may enhance memory if it promotes active learning on the part of the student. As always, there are individual differences among readers. Researchers have documented a number of counterintuitive effects that reflect various tradeoffs and differences among readers (Goldman and Saul, 1990; Mayer and Sims, 1994). For example, an outline that has a poor match to a text may promote active learning at the deep situation model level, but at the same time reduce memory for explicit text (McNamara et al., 1996;

Mannes, (1994). According to these findings, it is critical to specify the level of text represntation that is influenced by manipulations of auxiliary organizers.

One of our favorite gems of wisdom is that there is nothing more practical than a good theory. In this spirit, cognitive models of discourse can be used as a fine-tuned guide to revise text. For example, Britton and Gulgoz (1991) used Kintsch and van Dijk's (1978) model to revise expository texts. The Kintsch and van Dijk model identifies points in the text where there are coherence gaps. Readers are expected to generate inferences to fill these gaps but sometimes they fail and comprehension suffers. Britton and Gulgoz (1991) prepared revised versions of the texts that made these critical coherence-based inferences explicit. Memory for the revised texts was substantially superior to memory for the original texts with coherence gaps. Moreover, the model-based revisions were more memorable than revisions prepared by text linguists and by professional writers for major magazines (such as *Time* and *Life*).

Summary and Conclusions

Studies of cognition and discourse have dramatically increased during the short 25-year history of the field. Early research tested theories of discourse in sister fields, such as text linguistics, computational linguistics, and artificial intelligence. The early work focused on the structures and processes associated with the explicit text and on coherence at various levels. Contemporary research continues to test theories of discourse representation in other fields, but there is a more accurate sense of the plausible representations and processes in humans. There has been a shift toward understanding how these representations are shaped by world knowledge and comprehension strategies. Researchers hardly deny the importance of explicit text, but explicit text is only one piece of the puzzle.

Cognitive researchers have developed sophisticated models of discourse comprehension and production that attempt to explain complex patterns of experimental data. These models of discourse are integrated with general theories of cognition, not just discourse *per se*. Symbolic theories assume that knowledge is represented in the form of propositions, conceptual structures, schemata, and production rules. This content is dynamically activated and created in a limited-capacity working memory as comprehension proceeds. Connectionist theories assume that knowledge is represented in a more distributed mass of neural units, connected by excitatory or inhibitory weights. The activation values of the neural units dynamically change in working memory as comprehension proceeds. The two major models of discourse comprehension, the construction-integration model and the CAPS/reader models, are hybrids of the symbolic and connectionist traditions. These models attempt to account for complex interactions among multiple levels of representation. Future research needs to expand these

models by tackling longer texts, global patterns of coherence, a broader array of inferences, richer situation models, diverse reader goals, and pragmatic constraints between communication participants.

An adequate cognitive model accounts for the final meaning representations that get constructed and the process of constructing these representations on-line. According to the structure building framework, for example, the comprehender first builds a framework for a particular structure, then adds new relevant information to the structure, and shifts to a different structure when incoming information is irrelevant. Tests of the final meaning representations involve the collection of memory and judgment data after comprehension is finished. Tests of on-line comprehension mechanisms involve the collection of processing time data, such as reading times, gaze durations, and word naming latencies for test words that are interspersed in the text during comprehension.

Inferences are generated during the course of building the meaning representations on-line. Researchers have proposed several models that predict which classes of knowledge-based inferences are generated during comprehension. Available data appears to best fit a constructionist theory. This theory states readers construct those inferences that are relavet to the readers' comprehension goals, that establish local and global coherence, and that explain why propositions are mentioned in the text. Future research on inference generation needs to identify the precise conditions under which particular classes of inference are generated, as well as the time-course of their generation. There are almost no data on global thematic inferences and on inferences about the pragmatic communication between writer and reader. Future research needs to continue contrasting different discourse genres. A crude classification distinguishes narrative, expository, persuasive, and descriptive texts. A more mature classification scheme would include numerous subcategories, as well as hybrids. The representations, comprehension strategies, and pragmatic assumptions differ substantially among the various genres. One of the refreshing recent trends has been to focus on naturalistic literary works, such as novels and short stories. Studies of literary comprehension should help us understand the complex relationships among discourse, cognition and emotion.

Cognitive models of discourse have periodically been applied to real world problems. The models provide some guidance in revising text to make it more memorable or easy to comprehend. These efforts have proven useful in several practical arenas, such as education, the design of texts, surveys, and questionnaires, and the design of computer systems. We anticipate that cognitive research will have an important role in the future development of the 'information highway'.

Recommended Further Reading

Gernsbacher's *Handbook of Psycholinguistics* (1994) provides an excellent survey of contemporary research in the psychology of language and discourse. Introductory books by Just and Carpenter (1987) and Singer (1990) introduce the newcomer to cognitive research on language and discourse.

Cognitive psychologists have devoted considerable effort to building computational and mathematical models of discourse. Recent edited books provide a comprehensive survey of these sophisticated models (Britton and Graesser, 1995; Weaver et al., 1995). Original journal articles are available for those who wish to pursue the construction-integration model (Kintsch, 1988) and the CAPS/reader model (Just and Carpenter, 1992).

We have recommendations for those who wish to pursue particular problems in discourse and cognition. Gernsbacher (1990) discusses the structure building framework and supporting research. An edited book by Lorch and O'Brien (1995) presents psychological research on text coherence. The problem of inference generation is covered in an edited book by Graesser and Bower (1990) and in two journal articles (Graesser et al., 1994; McKoon and Ratcliff, 1992). A book of readings edited by Clark (1993) discusses psychological research on the use of language in conversation. For those interested in the comprehension of literature and figurative language, we recommend an edited book by Kreuz and MacNealy (1995). The *Handbook of Reading Research* (Barr et al., 1991) covers educational research on different discourse genres.

References

Albrecht, J.E. and O'Brien, E.J. (1993) 'Updating a mental model: maintaining both local and global coherence', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19: 1061–70.

Anderson, J.R. (1983) *The Architecture of Cognition.* Cambridge, MA: Harvard University Press.

Barr, R., Kamil, M.L., Mosenthall, P. and Pearson, P.D. (eds) (1991) *Handbook of Reading Research.* London: Longman.

Bates, E. (1976) *Language and Context: the Acquisition of Pragmatics.* New York: Academic Press.

Bower, G.H., Black, J.B. and Turner, T.J. (1979) 'Scripts in memory for text', *Cognitive Psychology,* 11: 177–220.

Brewer, W.F. and Ohtsuka, K. (1988) 'Story structure, characterization, just world organization, and reader affect in American and Hungarian short stories', *Poetics*, 17: 395–415.

Britton, B.K. and Graesser, A.C. (eds) (1995) *Models of Understanding Text.* Hillsdale, NJ: Erlbaum.

Britton, B.K. and Gulgoz, S. (1991) 'Using Kintsch's computational model to improve instructional text: effects of repairing inference calls on recall and cognitive structures', *Journal of Educational Psychology,* 83: 329–45.

Clark, H.H. (1993) *Arenas of Language Use.* Chicago: University of Chicago Press.

Clark, H.H. and Clark, E.V. (1977) *Psychology and Language.* New York: Harcourt Brace Jovanovich.

Clark, H.H. and Lucy, P. (1975) 'Understanding what is meant from what is said: a study in conversationally conveyed requests', *Journal of Verbal Learning and Verbal Behavior,* 14: 56–72.

Clark, H.H. and Schaefer, E.F. (1989) 'Contributing to discourse', *Cognitive Science,* 13: 259–94.

D'Andrade, R.G. and Wish, M. (1985) 'Speech act theory in quantitative research on interpersonal behavior', *Discourse Processes,* 8: 229–59.

Daneman, M. and Carpenter, P.A. (1980) 'Individual differences in working memory and reading', *Journal of Verbal Learning and Verbal Behavior,* 19: 450–66.

Dixon, P., Bortolussi, M., Twilley, L.C. and Leung, A. (1993) 'Literary processing and interpretation: towards empirical foundations', *Poetics,* 22: 5–34.

Dopkins, S., Klin, C. and Myers, J.L. (1993) 'Accessibility of information about goals during the processing of narrative texts', *Journal of Memory and Language,* 19: 70–80.

Fletcher, C.R. and Bloom, C.P. (1988) 'Causal reasoning in the comprehension of simple narrative texts', *Journal of Memory and Language,* 27: 236–44.

Fodor, J.D. (1983) *Modularity of Mind.* Cambridge, MA: MIT Press.

Frazier, L. and Fodor, J.D. (1978) 'The sausage machine: a new two-stage parsing model', *Cognition,* 6: 291–325.

Gernsbacher, M.A. (1990) *Language Comprehension as Structure Building.* Hillsdale, NJ: Lawrence Erlbaum.

Gernsbacher, M.A. (1993) 'Less skilled readers have less efficient supression mechanisms', *Psychological Science,* 4: 294–8.

Gernsbacher, M.A. (ed.) (1994) *Handbook of Psycholinguistics.* New York: Academic Press.

Gernsbacher, M.A., Goldsmith, H.H. and Robertson, R.R. (1992) 'Do readers mentally represent character's emotional states?', *Cognition and Emotion,* 6: 89–112.

Gerrig, R.J. (1993) *Experiencing Narrative Worlds.* New Haven, CT: Yale University Press.

Gibbs, R. (1994) *The Poetics of Mind: Figurative Thought, Language, and Understanding.* New York: Cambridge University Press.

Givón, T. (1993) 'Coherence in text, coherence in mind', *Pragmatics and Cognition,* 1: 171–227.

Glucksberg, S., Gildea, P. and Bookin, H.B. (1982) 'On understanding nonliteral speech: can people ignore metaphor?', *Journal of Verbal Learning and Verbal Behavior,* 21: 85–98.

Golden, R.M. and Rumelhart, D.E. (1993) 'A parallel distributed processing model of story comprehension and recall', *Discourse Processes,* 16: 203–37.

Goldman, S.R. and Saul, E.U. (1990) 'Flexibility in text processing: a strategy competition model', *Learning and Individual Differences,* 2: 181–219.

Goldman, S.R. and Varma, S. (1995) 'CAPing the construction-integration model of discourse comprehension, in C. Weaver, S. Mannes and C. Fletcher (eds), *Discourse Comprehension: Models of Processing Revisited.* Hillsdale, NJ: Erlbaum.

Graesser, A.C. (1981) *Prose Comprehension beyond the Word.* New York: Springer Verlag.

Graesser, A.C. and Bower, G.H. (eds) (1990) *The Psychology of Learning and Motivation: Inferences and Text Comprehension.* San Diego: Academic Press.

Graesser, A.C., Gordon, S.E. and Sawyer, J.D. (1979) 'Memory for typical and atypical actions in scripted activities: test of a script + tag hypothesis', *Journal of Verbal Learning and Verbal Behavior,* 18: 319–32.

Graesser, A.C., Singer, M. and Trabasso, T. (1994) 'Constructing inferences during narrative text comprehension', *Psychological Review,* 101: 371–95.

Grice, H.P. (1975) 'Logic and conversation', in P. Cole and J. L. Morgan (eds), *Syntax and Semantics. Vol. 3: Speech Acts.* New York: Academic Press. pp. 41–58.

Haberlandt, K. and Graesser, A.C. (1985) 'Component processes in text comprehension and some of their introductions', *Journal of Experimental Psychology: General,* 114: 357–74.

Haenggi, D., Kintsch, W. and Gernsbacher, M.A. (1995) 'Spatial models and text comprehension', *Discourse Processes,* 19: 173–99.

Halliday, M.A.K. and Hasan, R. (1976) *Cohesion in English.* London: Longman.

Haviland, S.E. and Clark, H.H. (1974) 'What's new? Acquiring new information as a process in comprehension', *Journal of Verbal Learning and Verbal Behavior,* 13: 512–21.

Hess, D.J., Foss, D.J. and Carroll, P. (1995) 'Effects of global and local context on lexical

processes during language comprehension', *Journal of Experimental Psychology: General*, 124: 62–82.
Just, M.A. and Carpenter, P.A. (1987) *The Psychology of Reading and Language Comprehension*. Boston, MA: Allyn and Bacon.
Just, M.A. and Carpenter, P.A. (1992) 'A capacity theory of comprehension: individual differences in working memory', *Psychological Review*, 99: 122–49.
Kintsch, W. (1974) *The Representation of Meaning in Memory*. Hillsdale, NJ: Erlbaum.
Kint ch, W. (1988) 'The role of knowledge in discourse comprehension: a constructive-integration model', *Psychological Review*, 95: 163–82.
Kintsch, W. and van Dijk, T.A. (1978) 'Toward a model of text comprehension and production', *Psychological Review*, 85: 363–94.
Kreuz, R.J. and MacNealy, M.S. (eds) (1995) *Empirical Approaches to Literature and Aesthetics*. Norwood, NJ: Ablex.
Lightman, A. (1993) *Einstein's Dreams*. New York: Warner.
Long, D.L., Golding, J.M. and Graesser, A.C. (1992) 'The generation of goal related inferences during narrative comprehension', *Journal of Memory and Language*, 5: 634–47.
Lorch, R. and O'Brien, E. (eds) (1995) *Sources of Coherence in Reading*. Hillsdale, NJ: Erlbaum.
Magliano, J.P., Baggett, W.B., Johnson, B.K. and Graesser, A.C. (1993) 'The time course of generat ng causal antecedent and causal consequence inferences', *Discourse Processes*, 16: 35–53.
Mandler, J.M. and Johnson, N.S. (1977) 'Remembrance of things parsed: story structure and recall', *Cognitive Psychology*, 9: 111–51.
Mannes, S.M. (1994) 'Strategic processing of text', *Journal of Educational Psychology*, 86: 577–88.
Mayer, R.E. and Sims, V.K. (1994) 'For whome is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning', *Journal of Educational Psychology*, 86: 389–401.
McClelland, J.L. and Rumelhart, D.E. (eds) (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 2. Cambridge, MA: MIT Press.
McKoon, G. and Ratcliff, R. (1992) 'Inference during reading', *Psychological Review*, 99: 440–66.
McNamara, D.S., Kintsch, E., Songer, N.S. and Kintsch, W. (1996) 'Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text', *Cognition and Instruction*, 14 (1): 1–43.
Meyer, B.F.J. (1975) *The Organization of Prose and its Effect on Memory*. New York: Elsevier.
Miall, D.S. and Kuiken, D. (1994) 'Beyond text theory: understanding literary response', *Discourse Processes*, 17: 137–52.
Millis, K. and Just, M.A. (1994) 'The influence of connectives on sentence comprehension', *Journal of Memory and Language*, 33: 128–47.
Morrow, D.G., Greenspan S.L. and Bower, G.H. (1987) 'Accessibility and situation models in narrative comprehension', *Journal of Memory and Language*, 26: 165–87.
Myers, J.L., Shinjo, M. and Duffy, S.A. (1987) 'Degree of causal relatedness and memory', *Journal of Memory and Language*, 26: 453–65.
Perfetti, C.A. (1990) 'The cooperative language processors: semantic influences in an autonomous syntax', in D.A. B lota, G.B. Flores d'Arcais and K. Rayner (eds), *Comprehension Processes in Reading*. Hillsdale, NJ: Erlbaum.
Potts, G.R., Keenan, J.M. and Golding, J.M. (1988) 'Assessing the occurrence of elaborative inferences: lexical decision versus naming', *Journal of Memory and Language*, 27: 399–415.
Rinck, M. and Bower, G.H. (1995) 'Anaphora resolution and the focus of attention in situation models', *Journal of Memory and Language*, 34: 110–31.
Schank, R.C. and Abelson, R.P. (1977) *Scripts, Plans, Goals, and Understanding: an Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Erlbaum.
Searle, J.R. (1969) *Speech Acts*. London: Cambridge University Press.
Singer, M. (1990) *Psychology of Language*. Hillsdale, NJ: Erlbaum.

Spilich, G.J., Vesonder, G.J., Chiesi, H.L. and Voss, J.F. (1979) 'Test processing of domain-related information for individuals with high and low domain knowledge', *Journal of Verbal Learning and Verbal Behavior*, 18: 275–90.
St John, M.F. (1992) 'The story Gestalt: a model of knowledge intensive processes in text comprehension', *Cognitive Science*, 16: 271–306.
Stein, N.L. and Glenn, C.G. (1979) 'An analysis of story comprehension in elementary school children', in R.O. Freedle (ed.), *New Directions in Discourse Processing*, vol. 2. Norwood, NJ: Ablex.
Trabasso, T. and Suh, S.Y. (1993) 'Using talk-aloud protocols to reveal inferences during comprehension of text', *Discourse Processes*, 16: 3–34.
van den Broek, S. and Lorch, R.F. (1993) 'Network representations of causal relations of causal relations in memory for narrative texts: evidence from primed recognition', *Discourse Processes*, 75–98.
van Dijk, T.A. (1972) *Some Aspects of Text Grammars*. The Hague: Mouton.
van Dijk, T.A. and Kintsch, W. (1983) *Strategies of Discourse Comprehension*. New York: Academic Press.
van Oostendorp, H. and Zwaan, R.A. (eds) (1994) *Naturalistic Text Comprehension*. Norwood, NJ: Ablex.
Weaver, C.A., Mannes, S. and Fletcher, C.R. (eds) (1995) *Discourse Comprehension: Strategies and Processing Revisited*. Hillsdale, NJ: Erlbaum.
Whitney, P., Ritchie, B.G. and Clark, M.B. (1991) 'Working memory capacity and the use of elaborative inferences in text comprehension', *Discourse Processes*, 14: 133–45.
Zwaan, R.A., Magl ano, J.P. and Graesser, A.C. (1995) 'Dimensions of situation model construction in narrative comprehension', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21: 386–97.