GERNSBACHER, M. A., & KASCHAK, M. (2013).
Text comprehension. In D. Reisberg (Ed.), The Oxford
Handbook of Cognitive Psychology (pp. 462-474).
London: Oxford University Press.

Text Comprehension

Morton Ann Gernsbacher and Michael P. Kaschak

Abstract

The study of text comprehension is the study of the cognitive processes involved as people process (and ultimately understand) the words, phrases, and sentences that make up larger bodies of language use (e.g., stories, magazine articles, novels, and so on). This chapter provides an introduction to several aspects of research on text comprehension: the methods commonly employed to study text comprehension, the major themes that have emerged over the past several decades of work in this field, and the theories that have been proposed to explain the comprehension process. In examining each of these aspects of the field, we highlight not only the state of the art in what is currently known about text comprehension but also the wide range of techniques and research questions that have come to characterize this area of psychological research.

Key Words: language comprehension, situation models, mental models, discourse

The use of language permeates our existence. We spend virtually all of our waking moments (and perhaps many of our sleeping moments) engaged in some kind of linguistic activity—writing e-mail, having conversations with friends, running through thoughts in our heads, and the like. A common factor across these situations is that language is almost always experienced in chunks larger than single words or single sentences. Thus, in most cases the comprehension of language requires not just understanding the meaning of the individual words or sentences, but the integration of the meaning of those words or sentences into a larger understanding of what is being talked or read about. The processes involved in comprehending these larger units of language (stories, newspaper articles, conversations, etc.) have been studied in some detail over the past several decades (see Graesser, Gernsbacher, & Goldman, 2003, for reviews). The present chapter provides an overview of research in this area, with a particular focus on reviewing the work that has examined comprehension of texts.

Text comprehension is a complex cognitive activity. To illustrate the complexity, consider what it takes to understand the following passage:

> Jane woke up in a panic on Thursday morning. Her rent check was due the next day, but there was no money in her bank account. When she walked into the kitchen, she remembered that she had not yet deposited the large refund check that she had gotten from the IRS. After getting dressed, she grabbed the check and drove to the bank.

At the most basic level, understanding this passage requires the reader to decode the orthographic symbols on the page, to recognize the words that the clumps of symbols form, and to recover the syntactic structure of the sentences. Beyond this, the reader must make sense of the whole text. The reader is told that Jane is in a panic, that her rent is due, and that she has no money. It seems clear that these statements are related (i.e., Jane is panicked because she cannot pay her rent), but the text does not state these relations explicitly; rather, readers must infer the relations based on their knowledge of the world. Jane's goal, namely to reduce her distress by finding a way to pay the rent, is also not explicitly stated in the text. The next sentence about her check from the IRS fits within the structure of this goal and therefore is easily incorporated into the ongoing situation. Finally, we are told that Jane drives to the bank as a way of meeting her goal. Note that if the final sentence had indicated that Jane drove to the beach, most readers would be confused, and their confusion would indicate a clash between their understanding of the situation conveyed by the text and their world knowledge: People who need to pay their rent do not usually take important checks to the beach.

This brief example highlights the major themes that have characterized research on text comprehension. Readers need to establish a representation to keep track of the events in the story, to be aware of the goals of the characters, and so on. Readers also need to draw on their knowledge of the world to draw inferences that fill the blanks of components of the situation not explicitly described. Finally, readers need to monitor the information that is presented in the text to make sure it is coherent with the representation that they have constructed based on the previous sentences. The main focus of this chapter will be to describe the cognitive mechanisms that underlie readers' ability to accomplish these tasks.

Methods of Studying Text Comprehension

Text comprehension is a complex activity that involves cognitive operations on every level of linguistic representation. Accordingly, a broad range of research methodologies has been used to investigate the mechanisms involved in the comprehension process. In this section of the chapter, we discuss the methods that are most frequently used in the field.

Verbal Protocols

When thinking about the question of how a text is understood, one might be interested in noting the kinds of overt strategies that readers employ to structure the comprehension task (Graesser, 2007). The collection of verbal protocols, or "think aloud" measures (see Ericsson & Simon, 1993), provides an efficient means of doing so. Research participants may be asked to provide a report of whatever they are thinking as they read a text, or they may be asked to respond to specific questions. For example, at a given point during a text, a reader may be asked why a character performed a particular action, what the reader thinks will happen next in the story, or how well the reader thinks she or he is understanding what she or he is reading. Although verbal reports may not accurately reflect the moment-by-moment processing that goes into the comprehension of a text, they do provide a valuable source of information about the ways that readers are approaching the comprehension task (e.g., what elements of the text they are paying attention to and which elements they are ignoring). Verbal reports can also be used in conjunction with other research methods (see Magliano & Graesser, 1991; Trabasso & Suh, 1993).

Probe Response Measures

Although verbal protocols provide researchers with insight into the knowledge that readers are retrieving as they process a text, verbal protocols do not allow researchers to ascertain when this knowledge is retrieved during the real-time, moment-by-moment processing involved in comprehension. A widespread method for accomplishing that goal is the use of probe response tasks. In this method, research participants are asked to read a text. At various points in their reading, they are interrupted with a secondary task that requires them to respond to a single word. For example, participants may read a sentence such as (1).

(1) John walked down the street with a smile on his face.

After reading this sentence, the word *happy* may be presented visually or auditorily. The reader may be asked to decide whether *happy* is a word (i.e., a *lexical decision task*) or may be asked to simply say the word *happy* aloud (i.e., a *naming* task). The logic of these tasks is that if the reader has used the information presented in sentence (1) to infer that John is happy, then the response to the probe task should be faster when the probe word is *happy* (or, a related word, such as *joyful*) than when the probe word is *sad* (or a word completely unrelated to this inference, e.g., *table*).

Reading Time Measures

Whereas probe response tasks have provided a valuable tool for studying the moment-by-moment retrieval of knowledge during text comprehension, they have the unfortunate feature of disrupting the normal reading process: To respond to the probe, participants need to stop reading and attend to a secondary task. Therefore, researchers have used other methods to assess the reading process under more normal processing conditions. Reading time

methods range from those that are fairly coarse (e.g., measuring the time it takes readers to process a whole clause or sentence) to those that are more precise (e.g., measuring the reading time for each word as the sentence is read, either by asking the reader to push a button to move from one word to the next in a sentence, or by monitoring the position of the eyes as they move across the sentence). Reading time measures are typically used as an index of processing difficulty. Increases in reading time suggest an increase in processing difficulty, which itself can be caused by many factors (e.g., the need to generate an inference to maintain the coherence of the text).

Brain Measures

Several methods of studying brain activity have been employed to study text comprehension. One such method involves measuring event related potentials (ERPs), which are changes in electrical activity in the brain that occur during text comprehension. For example, readers may be presented with a sentence such as (2).

(2) Jane wanted to deposit her paycheck, so she drove to the beach.

The word *beach* is anomalous in this context, and so it is expected that a change in brain activity should occur as participants read the word and notice the inconsistency. ERP methods provide very good temporal resolution in that the changes in brain activity are closely time-locked to particular events that occur while processing language. However, because ERPs are detected by electrodes placed on the scalp, ERPs are not very good for localizing the activity to a particular part of the brain. Other measures of brain activity are functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). Both fMRI and PET use changes in blood flow in the brain to determine which brain regions are involved in a given cognitive task. These methods are useful for localizing changes in brain activity that occur during text comprehension, but they suffer from relatively poor temporal resolution due to the slow rate of onset and offset for changes in blood flow in the brain, which is unfortunate considering how rapidly a process like text comprehension occurs.

Theoretical Issues in Text Comprehension

Our discussion of the processes involved in text comprehension will be centered on several theoretical issues that have received attention in the literature. The first issue concerns the kinds of representation that are constructed during text comprehension. Here we will consider the multiple ways that a text can be represented and the role that each of these representations can play in the comprehension process. The second issue concerns the medium that is used to construct representations of text content. The final issue we tackle concerns the processes involved in comprehending a text. Of concern here will be the moment-by-moment processes involved in text comprehension, particularly the ways that readers "fill in the blanks" so that they can understand what the text is trying to convey, even though the text itself may leave important information unsaid.

Levels of Representation

Beginning with van Dijk and Kintsch (1983), approaches to text comprehension have broadly accepted the claim that readers construct three levels of representation during the comprehension process. The first of these is the *surface-level* representation, which is a verbatim representation of the wording used in the text. Generally speaking, the surface-level representation is short lived (e.g., Barclay, Bransford, Franks, McCarell, & Nitsch, 1974; Gernsbacher, 1985). Kintsch et al. (1990) asked participants to read texts, then gave them a recognition test to assess their memory for the different levels of text representation. Whereas memory for the surface form declined quickly (and was almost entirely lost over a period of a few days), memory for textbase and situation model representations remained somewhat accurate even several days after encountering the original texts. Nonetheless, there are cases in which memory for the surface form of the text is long-lasting. For example, Murphy and Shapiro (1994) demonstrated that memory for the surface form of a text can be high in cases where the text presents a joke or an insult (both being contexts in which remembering the surface form of the language would be of some import).

The second layer of representation is the *textbase*. The textbase is a propositional representation of the ideas explicitly stated in the text. As seen in the brief story presented at the beginning of this chapter, texts are rarely explicit about everything that is relevant to the comprehension of the situation that is being described. The textbase level of representation does not function to fill in all of these gaps, but it does begin to glue some of the elements presented in the text together. Consider sentence (3).

(3) The boy gave his mother a cupcake.

A verb such as *give* has a particular argument structure, or configuration of noun phrases, that fills different roles in the action of giving. The boy is referred to as the *agent* of the action (i.e., the "doer"), the mother is the *recipient* of the action (i.e., the "receiver"), and the cupcake is the *patient* of the action (i.e., the thing being acted on). The binding of elements in the text (boy, mother, and cupcake) to particular roles in the action *(agent, recipient, patient)* is accomplished in the textbase level of representation. It is worth noting that although the textbase is traditionally considered to be a level of representation in discourse processing, there are some who argue that this level of representation is unnecessary. For example, it has been proposed that individual lexical items may provide the comprehension system with "instructions" about how to produce the situation model, thus eliminating the need to have a textbase that connects the surface level and situation model levels of representation (e.g., Gernsbacher, 1990).

The third layer of representation constructed during text comprehension, and the one that has received the most attention, has been variously called the *mental model, discourse model,* or *situation model.* These terms have been used for very specific purposes in the literature, but for the present purposes we will treat them as equivalent and adopt the convention of referring to this level of representation as the situation model. The situation model is a representation of what the text is about. It is an amalgam of information that is contained in the textbase and information that has been retrieved from the comprehender's general store of world knowledge (Kintsch, 1988). It is here that information not explicitly mentioned in the text can be brought to bear in the service of understanding the events that are being described. The generation of the inferences necessary to make sense of the content of the text takes place during the construction of a situation model.

Zwaan et al. (1995; Zwaan & Radvansky, 1998) proposed that situation models are structured around five dimensions:

Protagonist: Who is involved in the events being described?
Time: At what time is the event taking place?
Space: What are the spatial relations between the characters, objects, and events being described?
Causality: Which events in the text are causally related to each other?
Intentionality: Is the incoming information relevant to the protagonist's goals and intentions?

Readers track information along these, and possibly other, dimensions as they process a text. Zwaan et al.'s (1995) account has two important implications. First, events that are interrelated among several dimensions will be more closely linked in memory than events that are interrelated among fewer dimensions. Second, the ease with which incoming information will be integrated into the existing situation model depends on the number of dimensions on which the incoming text content matches the immediately preceding content. Overall, the more overlap between the incoming content and the existing state of the situation model, the easier it will be to integrate the new information into the model.

A number of studies support Zwaan et al.'s (1995) general claim that readers monitor several dimensions during the comprehension of text. Space is the most explored of the dimensions. Some early views of situation models held that comprehenders constructed spatial representations of text content in a spatial medium that was analogous to the three-dimensional world in which we live (Glenberg, Kruley & Langston, 1994; Johnson-Laird, 1983). Rinck and Bower (1995; see also Morrow, Bower & Greenspan, 1989) found evidence for this claim in experiments in which participants were asked to read texts about the protagonists' movements around a building. They found that the reading time for sentences referring to objects in the building increased as a function of the number of rooms between the location of the object and the current location of the main character in the building. That is, the more "space" the participants needed to cover to get from their current location to the location of the object that was mentioned, the harder it was to process the sentence.

Early indications were that situation models contained analogical spatial representations; however, subsequent work put important qualifications on this claim. Langston, Kramer, and Glenberg (1998) report a series of studies in which they failed to find anything like analogical spatial representation. Moreover, Radvansky and Copeland (2000) used a combination of reading time measures and memory assessments to demonstrate that sentences describing functional spatial relationships between entities are read faster and remembered better than sentences describing nonfunctional relationships. Current thinking on the role of space in situation models has thus shifted somewhat from the initial positions that appeared in the literature. Specifically, it appears that readers do not routinely construct detailed spatial

representations in their situation models unless (a) the text is explicitly spatial, or (b) there is some strategic reason to do so (see Langston et al., 1998, for a discussion). Furthermore, the spatial relationships that are constructed are rarely analogical.

Temporal representations in situation models have also been studied. Zwaan (1996) reports a seminal set of experiments that explored the processing of sentences that indicate a time shift in narratives, as in (4).

(4a) A moment later, the fireman....
(4b) A day later, the fireman...
(4c) A month later, the fireman...

Zwaan (1996) found that the processing speed of such sentences was related to the magnitude of the time shift that was indicated. The longer the time shift, the longer the processing time. This seems to indicate that comprehenders are sensitive to the magnitude of time shifts as they construct their situation models—longer time shifts represent larger breaks in the timeline of the narrative and therefore require more updating of the situation.

The other dimensions proposed by Zwaan et al. (1995) have received comparatively less attention in the literature (see Zwaan & Radvansky, 1998, for a review). Nonetheless, there is empirical support for the general notion that these dimensions of a situation are tracked during text comprehension. It is important to note that most studies of the dimensional approach to text processing examine the tracking of a single dimension; few studies have explored multiple dimensions at once (but see Therriault, Rinck, & Zwaan, 2006, for an example). Gaining an understanding of how comprehenders simultaneously track multiple dimensions during text comprehension will be an important topic for further research in this area.

What Are Situation Models Made of?

The situation model is the level of representation that is taken to be synonymous with full comprehension of a text. It is here that comprehenders integrate information presented in the text with their general store of knowledge and use this integration to fashion an understanding of the situations that are described. While virtually all language comprehension researchers would agree that a situation model (or something like it) is essential for successful text comprehension, there has been somewhat less agreement about the nature of situation models themselves.

One prominent view of the representational medium of situation models has been to propose that the models are, like the textbase, propositional in nature (Kintsch, 1988, 1998). To illustrate this type of representation, consider again sentence (3): The boy gave his mother a cupcake. The nature of this action could be represented via the proposition (GIVE, AGENT:boy, PATIENT:cupcake, RECIPIENT:mother). This view holds that knowledge (generally speaking) is stored in propositional format, that propositions can be derived explicitly from the text when forming the textbase, and these propositions can be united with propositions from the reader's general knowledge base to form a situation model. Proposition-based accounts of situation models have achieved wide-ranging success in explaining empirical phenomena associated with situation models. As one example, the activation and inhibition of propositions, as instantiated in Kintsch's (1988) Construction-Integration model, has proven to be relatively successful in predicting what information will be available (and what information will not be available) at various points during the comprehension of a text, the kinds of inferences that readers will make as reading a text, and the kinds of performance that students will demonstrate as they comprehend and solve mathematical word problems (a task that inherently involves the comprehension of text).

An alternative view of situation models holds that the representational medium is not propositional, but rather analogical. The key distinction between these views is that whereas the propositional approach holds that the information presented in the text is converted into a propositional "language of thought," the analogical view holds that the information presented in the text is used to retrieve representations that are akin to perceptual and spatial representations of the sort that forms our mental imagery (putting aside the contentious debate about whether imagery comes from propositional representations). Johnson-Laird's (1983) approach to mental models is an early example of this perspective. A more recent exemplar of this view is the claim that language comprehension requires the construction of sensorimotor simulations of the events being described (e.g., Barsalou, 1999; Zwaan, 2004). That is, a sentence such as "The car approached you" would be understood by using the visual system to simulate what it would look like for a car to come toward you (Kaschak et al., 2005), and a sentence such as "The boy gave his mother a cupcake" would be understood by using the motor system to simulate the action of transferring something from one person to another (Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006).

Both propositional and analogical representations have their proponents in the literature. Proponents of the propositional representation format point to the empirical successes of the formal models that have been constructed based on this sort of representation. As noted earlier, Kintsch's (1988) Construction-Integration theory has been employed to model a range of comprehension data in the text-reading literature and in educational settings. More broadly, there is a large body of literature showing that nonanalogical representations play a role in language processing. Examples of such representations include knowledge of abstract syntactic rules (e.g., Frazier & Clifton, 1996) and probabilistic knowledge concerning the patterns of use of individual words, sentence constructions, syllables, and the like (e.g., Landauer & Dumais, 1997; MacDonald, Pearlmutter, & Seidenberg, 1994; Roland, Dick, & Elman, 2007). This sort of information has been shown to play a role in language acquisition (e.g., Saffran, Aslin, & Newport, 1996), language comprehension (e.g., MacDonald et al., 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998), and language production (e.g., Chang, Dell, & Bock, 2006).

Proponents of analogical representations point to a growing body of evidence from behavioral studies (e.g., Glenberg & Kaschak, 2002; Meteyard, Bahrami, & Vigliocco, 2007; Zwaan & Taylor, 2006) and neuroimaging studies (e.g., Buccino et al., 2005; Glenberg et al., 2008; Hauk, Johnschrude, & Pulvermuller, 2004; Speer, Reynolds, Swallow, & Zacks, 2009) showing that the systems involved in perception and action planning are recruited during the comprehension of sentences describing both concrete and abstract situations. Behavioral studies have shown that the comprehension of sentences about action (e.g., "Open the drawer") facilitates the preparation and execution of congruent actions (e.g., pulling your arm toward your body; Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006), and both that the comprehension of language about motion affects one's acuity in perceiving motion in a visual stimulus (e.g., Meteyard et al., 2007) and that the perception of visual motion affects the comprehension of sentences about motion (e.g., Kaschak et al., 2005). Consistent with these behavioral results, fMRI studies have shown that asking participants to process words about perception and action activates the neural regions that have been identified as being involved in the processing of perceptual or motor information. As one example, Hauk et al. (2004) demonstrated that the parts of the motor strip that have been identified as controlling the lips, fingers, and legs were activated when participants processed the words lick, pick, and kick, respectively.

Given that situation models are central to the process of comprehending language, a resolution of the question about the representational format of such models will be important to developing an understanding of what it means to understand. It is not our goal to resolve this issue here, but a few points can be made. First, most studies of language comprehension are not incisive on the matter of what kinds of representation underlie the comprehension process. With the exception of some recent studies of the interplay between language comprehension and systems of perception and action planning, task performance in studies of language comprehension are generally open to explanation from both a propositional and analogical perspective (witness the parallels between this debate and the analogical/propositional debate in the literature on mental imagery; e.g., Kosslyn, 1994). More studies need to assess the nitty-gritty of the representational formats that are used in text comprehension.

Second, it is possible that some types of abstract representational formats (e.g., some types of propositional representations) may turn out to be represented analogically. As one example of this, Lakoff and Nunez (2000) argue that abstract mathematical representations (such as mathematical sets) can be represented in terms of concrete, analogical representations (e.g., a set can be represented as a container). Third, although some have called for approaches that incorporate both abstract types of linguistic representation and analogical, experiential types of representation (e.g., Louwerse, 2008), it remains to be seen whether such hybrid approaches can produce a satisfactory theory of language comprehension.

Processing a Text

At the beginning of this chapter, we walked through a brief example of what would be involved in comprehending a short text about a person who needed to pay her rent. It was clear that understanding even a short, relatively straightforward text involves a host of cognitive processes ranging from the detection and recognition of letters to the construction of inferences needed to fill in the gaps in the information presented explicitly in the text. In

this section of the chapter, we discuss the major processes involved in comprehending a text.

FIRST THINGS FIRST

Although the processing of all sentences in a text requires some degree of cognitive work, a number of studies have shown that the processing of initial sentences in a text requires more effort than the processing of subsequent sentences. Gernsbacher (1990) has described this phenomenon in terms of *laying a foundation* for subsequent comprehension to take place. Thus, processing the first sentences of a text involves the initial construction of a situation model, and this initial construction involves extra processing effort since the model needs to be started from scratch. Once the model is constructed, it becomes easier to add more information as the text proceeds. It should be noted that this phenomenon also holds for segments of a text where shifts of various sorts occur. For example, a comprehender may create a new submodel within a larger situation model of a text to represent an extensive flashback that is presented. The creation of this submodel requires additional processing work and thus results in slower comprehension of the first few sentences after the shift takes place.

GATHERING RAW MATERIALS

The information presented explicitly in a text forms the basis of text comprehension. The words and phrases that are used provide the initial triggers for retrieving information from memory, and it is the combination of information retrieved directly based on the words and phrases used in the text and information that is already present in the reader's situation model that combine to form the interpretation of a given sentence.

We are describing the process of using lexical information to retrieve information from memory as a process of gathering raw materials for the construction of a situation model largely because this initial stage of processing is initially indiscriminate with respect to what is retrieved. Beginning with Swinney (1979), it has been shown that a wide range of information is activated when one initially processes a word. Consider sentence (5).

(5) The spy placed a bug on the table.

Swinney (1979) asked participants to listen to sentences such as (5). When the critical word (in this case, *bug*) was reached, participants were visually presented with a string of letters on a computer screen and participants had to decide whether it was a word. The words that were presented were either related to the different senses of the critical word (e.g., *spy* and *insect* are related to different meanings of *bug*) or were unrelated to the critical word (e.g., *sew*). Swinney (1979) found that when the test words were presented immediately after the critical word in the sentence, participants were fast to respond to the words related to the meaning of the critical word and comparatively slower to respond to the unrelated word. This suggests that comprehenders immediately access all relevant information when processing a word with multiple meanings. Nonetheless, it has also been shown that the contextually inappropriate meanings of the word are quickly dampened; activation of inappropriate meanings seems to decay within around 750 msec from the processing of the word (Gernsbacher & Faust, 1995).

Successful language comprehension requires the resolution of ambiguity. Ambiguity exists at every representational level of language (phonology, lexicon, syntax, etc.), and most theories of language comprehension employ some mechanism through which ambiguity (e.g., selecting the meaning of an ambiguous word such as *bug*) is handled. For example, Gernsbacher (1990) suggests that ambiguities in word meaning are resolved by higher level representations inhibiting the contextually inappropriate lexical representations of the word in question (e.g., inhibiting the *insect* meaning of *bug* in a context about espionage). The quick resolution of ambiguity is central to comprehension as a failure to resolve ambiguities (e.g., by not selecting a single meaning of *bug*) can lead to the development of an incorrect representation of the content of the text. Indeed, Gernsbacher and colleagues have shown that an important distinction between good and poor language comprehenders is the ability to inhibit contextually inappropriate meanings. Gernsbacher and Faust (1995) review several studies using methods similar to the one employed by Swinney (1979) in which it is shown that whereas all readers initially activate a wide range of information relevant to the interpretation of a word with multiple meanings (such as *bug*), good comprehenders inhibit contextually inappropriate information within a few hundred milliseconds, but poor comprehenders do not. Thus, all comprehenders successfully recruit information for potential inclusion in one's representation of the meaning of a sentence or text, but poor comprehenders struggle because they are unable to successfully inhibit the information that is not relevant to the comprehension of the ongoing text.

ADDING TO THE STRUCTURE

Theories of text comprehension typically posit that information is added to a situation model on a clause-by-clause or sentence-by-sentence basis (e.g., Kintsch, 1988). Here, the information that is retrieved based on the explicit content of the text is integrated with both the existing structure of the situation model and additional knowledge that has been retrieved from memory. This process is referred to as *updating* the situation model. It is the updating of situation models that allows the comprehender to keep the content of a text together in a coherent representation.

The updating of situation models raises three issues. First, how does the reader know that a series of sentences is intended to be integrated into a single representation? There are several linguistic cues that signal the reader to interpret sentences as a discourse. Robertson et al. (2000) asked readers to process a series of sentences such as "The family rode together in a car." When the series of sentences began with the definite article *the*, readers tended to interpret the series of sentences as part of a connected discourse. On the other hand, when the same sentences were presented with the indefinite article *a*, readers tended to interpret the sentences as being unrelated to each other. The use of the article *the* implies that the entities being discussed are entities to which reference has already been made in the preceding sentences, and thus it cues the reader to treat the sentences as being related. Connectives such as *because*, *however*, *meanwhile*, and others play a similar role. For example, consider the sentences, "The grandparents prepared the food for their party. Meanwhile, the family was loading into their minivan." The focus of the two sentences is completely different, but the use of the connective *meanwhile* tells the reader that the two actions should be treated as part of the same event.

The second updating-related issue concerns the question of what information should be integrated into the situation model. Here, the linguistic distinction between *given* and *new* information (e.g., Haviland & Clark, 1974) is relevant. Given information is information that is already present in the situation model. New information is information that is being introduced in the current clause or sentence. The given information in the current sentence provides an anchor to the current state of the situation model, and it indicates which elements of the model will be updated. For example, mentioning the name of one character in a story will anchor the current sentence to the representation of that character in the situation model, and it will signal that the new information in the sentence can be added to the model to update the representations involving that character. Although much information is retrieved from memory during the processing of a clause, only that information that can be anchored to existing elements of the situation model will be easily integrated into the updated version of the model.

The final updating-related issue to be addressed here regards the means through which readers keep track of entities that are mentioned repeatedly in a text. When texts are written, they do not repeatedly use the same word to refer to a particular object or person. A character may be introduced as "Jane," and in the next sentence be referred to as "she," and later in the text be referred to as "the woman with the brown hair." To maintain order in his or her situation model, the reader must realize that "Jane," "she," and "the woman with the brown hair" all refer to the same person.

The most studied case of reference in discourse processing is the comprehension of *anaphoric* reference, or reference to a person or object that has previously been mentioned (as in the earlier example). One factor that influences the comprehension of anaphoric references is the nature of the information currently active in memory (and foregrounded in the mental model). If there is one female character currently in the foreground of the model (Jane), then pronouns such as "she" will be mapped onto Jane. If there is more than one female character in the foreground of the model, the pronoun will be mapped to the one whose representation is most active. In some cases, further knowledge is needed in order to determine the referent of an anaphor. When Jane is referred to as "the woman with brown hair," the reader needs to know that Jane has brown hair in order to easily understand the reference as intended.

The choice of reference for an entity in the text depends largely on the circumstances in which the entity is being mentioned. When a person or object is initially mentioned, it is usually marked with the article *a* or *an*, and the description provided is typically somewhat detailed. Subsequent references are less detailed and may be made by pronouns or other "shorthand" referring expressions. The more the entity has been mentioned in the text, the shorter the referring terms tend to be. These linguistic cues provide the reader with information as to whether a person or object that is being mentioned is one that has previously been mentioned in the text, or

whether a new entity has been introduced to the situation.

MAINTAINING COHERENCE

The preceding paragraphs have discussed several means through which readers integrate a series of sentences into a larger representation of the text. As seen in the examples that were provided, it is possible to create such larger representations by connecting information explicitly presented in the text to information that is already present in one's situation model. Such connections are made through processes of memory retrieval, where elements of the incoming language (e.g., a pronoun such as *she*, or a proper name such as *Jane*) serve as cues to retrieve information from the existing situation model (e.g., Myers & O'Brien, 1998). Thus, a pronoun such as *she* serves as a cue to retrieve the identity of a female person who is currently represented in the situation model. These memory-driven processes are no doubt an important component of text comprehension, but they are not sufficient to explain the entirety of how readers maintain a coherent representation of a text. To make sense of a text, readers often need to do additional processing work to figure out how the incoming sentence relates to the existing situation model. That is, they need to do processing work to maintain the coherence of the text as new information comes in.

Text comprehenders may strive to maintain coherence in their representation of a text in many ways. They might connect each incoming sentence to the most recently encountered sentences (as discussed earlier). In doing so, the reader is attempting to maintain *local coherence* (e.g., McKoon & Ratcliff, 1992) in his or her situation model. On the other hand, readers may attempt to maintain *global coherence* in their discourse model (e.g., Graesser, Singer, & Trabasso, 1994). That is, readers connect each incoming sentence to both the local content of the situation model and the global structure of the model (which includes information about the structure of the text, the overarching goals of the characters, and so on).

In general, readers attempt to maintain both local and global coherence when they construct mental models (e.g., O'Brien & Albrecht, 1992). Consider the short text presented at the beginning of this chapter. The narrative begins by presenting a goal for Jane (she needs to get money into her bank account). Several sentences later, the reader is told that Jane is going to the bank to deposit her check. As noted in our discussion of that example, if readers had been told that Jane is going to the beach with her check, they would no doubt find the sentence incongruous. Part of the reason that readers find that statement to be inconsistent with the rest of the text is that the behavior it describes is inconsistent with Jane's goals. That is, readers are keeping track of Jane's goals in order to maintain the global coherence of the text. In addition to tracking information needed to maintain global coherence, it is clear that readers attempt to maintain coherence from sentence to sentence (as seen in the previous section of this chapter; McKoon & Ratcliff, 1992). If readers fail to maintain global coherence in their discourse models, they may focus instead on maintaining local coherence. If a given text has too little local coherence, it will be viewed as altogether incoherent.

To achieve global and local coherence, the reader must often fill in details that are not explicitly presented in the text. That is, the reader must generate inferences about the events being described in order to maintain a coherent representation of the text. For example, when they encounter a pair of sentences such as "It was cold that morning. Joe slipped on the sidewalk," readers need to draw the inference that there may have been snow or ice on the sidewalk in order to integrate the two statements into a coherent discourse model. Although theories differ on the exact mechanisms through which inferences are generated, the general picture that emerges from the literature is that inference generation arises through the knowledge activation mechanisms that were discussed earlier in the context of "gathering raw materials" for the comprehension process. Comprehenders retrieve not just information relevant to the interpretation of individual words but also information associated with those meanings. For example, the *spy* meaning of *bug* implies not only espionage but also that the spy and the person being spied upon are enemies (or at least working for different government agencies). This knowledge forms the basis for inference generation, such that when we are told that one character in a story plants a bug in the hotel room of another character, we can use the knowledge retrieved to integrate the fact that the characters work for different governments into our situation model.

One of the major research questions about inference generation during text processing has centered on the issue of when inferences are generated and when they are not. According to the *minimalist* position (McKoon & Ratcliff, 1992), readers mostly attempt to maintain local coherence when they process a text. The only inferences that are routinely generated by readers are those that are required to maintain local coherence (such as the inference that there was ice on the sidewalk from the example in the previous paragraph). Readers are capable of drawing more global inferences from text, but these inferences are only drawn under certain circumstances (e.g., when the reader is attempting to process the text more deeply than usual). In contrast, the *promiscuous generation* position maintains that readers routinely generate a wide range of inferences from the text, including those that are not strictly necessary to ensure local coherence (Kintsch, 1988). These inferences include inferences about the goals of the characters, the emotional state of the characters, the cause-and-effect relationship between events in the text, the intent of the writer in conveying particular pieces of information, and so on.

A compromise between these extreme positions is the *constructivist* position (Graesser et al., 1994). The constructivist position holds that readers routinely draw inferences that meet their goals as comprehenders, inferences that maintain the coherence of the text, and inferences that explain why different events in the text are taking place. On this view, readers may appear to behave in accord with the minimalist position under certain conditions, and behave in accord with the promiscuous generation position in other conditions depending on the nature of the text and the goals they have in comprehending that text. For example, if the reader is attempting to skim the text in an effort to quickly glean information, he or she may draw few inferences (in keeping with the minimalist position). On the other hand, if individuals are reading a text for enjoyment (say, if they are reading a detective novel) or if they are trying to learn about a new field of study, they may read the text more carefully and draw a wider range of inferences (see Foertsch & Gernsbacher, 1994).

STRATEGIC PROCESSING

We conclude our discussion of the theoretical issues surrounding text comprehension by saying a few words about the ways that reading strategies play a role in the comprehension process. We have already discussed some of this in the context of constructionist approaches to text comprehension, where the general idea is that readers can choose to read a text in great detail or can choose to skim over the text and skip over a lot of the detail. Indeed, it has been noted that readers typically adopt a minimalist approach to comprehension and construct text representations that are just "good enough" to allow comprehension to proceed (e.g., Foertsch & Gernsbacher, 1994; Ferreira, 2003). It is clear that the amount of effort that a reader puts into comprehending a text has a strong influence on the representations that are constructed.

Another way that strategic processing can affect text comprehension is through the expectations that readers bring to bear about the text based on genre information. Experienced readers know that newspaper stories, detective novels, science fiction stories, and romantic comedies have typical structures, and knowledge of these structures can affect the ways that readers glean information from the text and the kinds of situation models that are constructed (e.g., Graesser, Kassler, Kreuz, & McLain-Allen, 1998; Zwaan, 1994).

Theories of Text Comprehension

In this section, we briefly introduce several theories that have been put forth to explain text comprehension.

Construction-Integration Model

Kintsch's (1988) Construction-Integration (CI) model is one of the most successful models of text comprehension. Although the model is not universally endorsed, there is a fairly wide consensus that the CI model captures nearly all of the basic processes that are required for successful comprehension. The CI model proposes that text is comprehended in two iterative stages. During the Construction stage, the incoming textbase enters working memory and retrieves potentially relevant information from long-term memory. This stage of processing happens quickly and automatically. During the Integration stage, the comprehension system assimilates the new information with the previously existing model of the discourse. The Integration stage is comparatively slow and resource consuming, as the comprehension system pares down the information activated in the Construction stage and integrates only the information that is most relevant to the present situation into the model of the discourse. The CI model provides a framework through which to understand the processes involved as information is activated, inhibited, and integrated into a coherent situation model, and as such can be applied to many of the research findings discussed throughout this chapter.

Structure Building Framework

The Structure Building Framework was outlined by Gernsbacher (1990). The theory proposes that discourse comprehension proceeds by building structured representations of the information presented in the textbase. The "structure" is based around the initial elements presented in the textbase. For example, if a narrative began, "Jane woke up on Thursday morning…," Jane would be the focus of the new structure. The process of initiating a structure is called *laying a foundation*. When new information is presented, it can either be *mapped* to the existing structure, or it can prompt the comprehension system to *shift* to a new structure. Finally, some information in the structure can be *enhanced* and made more available for further processing, whereas other information can be *suppressed* and made temporarily unavailable for further processing. The Structure Building Framework has been an influential account of the processes through which coherence is maintained in the ongoing comprehension of a situation model.

The Memory-Based Approach to Discourse Processing

Myers and O'Brien (1998) describe the memory-based approach to discourse processing. Whereas the Construction-Integration model and the Structure Building Framework posit the operation of "active" processing mechanisms (in the sense that these mechanisms actively retrieve or inhibit information when building a representation of the discourse), the memory-based approach is built on passive mechanisms of memory retrieval. The memory-based approach is based on Hintzman's (1986) MINERVA and its resonance process of memory retrieval. On this view, incoming information resonates both with the existing model of the discourse and with information in long-term memory. Information from memory is used to interpret the incoming sentence to the extent that it resonates with the new information. The primary contribution of the memory-based approach has been to outline a passive memory retrieval mechanism that serves as a theoretical alternative to the more active "activation and inhibition" mechanisms that have played a large role in many extant theories of text comprehension (e.g., Gernsbacher, 1990; Kintsch, 1988).

Event-Indexing Model

Zwaan et al.'s (1995) Event-Indexing model proposes that readers continually monitor the discourse model to maintain coherence on five dimensions: protagonist (who is involved in the events being described), time (when the events take place), space (where the characters, objects, and events are located), causality (why the events happen), and intentionality (what drives the protagonist, i.e., his goals and intentions). Shifts on any one of those dimensions—for example, if the text indicates a temporal delay between one event and the next—are typically associated with some processing cost as readers update their mental model. The largest processing costs tend to be associated with discontinuities on the time and protagonist dimensions. The Event-Indexing model has been influential in shaping research on how the dynamics involved in updating many aspects of a situation model at once affect the comprehension process.

Conclusions

Text comprehension is a complex cognitive operation that requires several levels of linguistic processing and the integration of information presented in the text with knowledge from the reader's long-term memory. As discussed in this chapter, research conducted during the past decades has productively enumerated the processes involved in text comprehension, and several valuable theories have been developed. Nonetheless, there are many exciting frontiers of text comprehension research that are only beginning to be explored. For example, several researchers have begun to apply basic research on text comprhension to problems related to the remediation of reading difficulties (e.g., Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007; see also chapters in McNamara, 2007). There is promise that research of the sort that has been reviewed in this chapter can contribute to the development of educational practices and interventions. As another example, there is growing interest in the question of whether the comprehension of texts reflects the same sort of processes that are involved in the comprehension of the events that occur in our lives. This research has suggested that similar principles can applied to the understanding of text processing and both event processing (e.g., Speer & Zacks, 2005) and the processing of episodes in films (e.g., Magliano, Miller, & Zwaan, 2001; Zacks, Speer, & Reynolds, 2009). Finally, the use of neuroimaging techniques such as fMRI to explore text comprehension promises to open new possibilities for exploring the processes and representations involved in the comprehension process (e.g., Robertson et al., 2000; Speer et al., 2009). Exploration of these new frontiers of research will no doubt deepen our understanding of text comprehension and ensure that the next decades of research in this area are as exciting and productive as the previous decades have been.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior, 13*, 471–481.

Buccino, G., Riggio, L., Hellia, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research, 24*, 355–363.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*, 2, 234–272.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (rev. ed.). Cambridge, MA: MIT Press.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203.

Foertsch, J., & Gernsbacher, M. A. (1994). In search of complete comprehension: Getting "minimalists" to work. *Discourse Processes, 18*, 271–296.

Frazier, L., &Clifton, C., Jr. (1996). *Construal*. Cambridge, MA: MIT Press.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17*, 324–363.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Faust, M. (1995). Skilled suppression. In F. N. Dempster & C. N. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 295–327). San Diego, CA: Academic Press.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review, 9*, 558–565.

Glenberg, A. M., Kruley, P., & Langston, W. E. (1994). Analogical processes in comprehension: Simulation of a mental model. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. San Diego, CA: Academic Press.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61*, 905–919.

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3–26). Mahwah, NJ: Erlbaum.

Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (Eds.). (2003). *Handbook of discourse processes*. Mahwah, NJ Erlbaum.

Graesser, A. C., Kassler, M. A., Kreuz, R. J., & McLain-Allen, B. (1998). Verification of statements about story worlds that deviate from normal conceptions of time: What is true about Einstein's Dreams? *Cognitive Psychology, 35*, 246–301.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371–395.

Hauk, O., Johnschrude, I., &Pulvermuller, F. (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301–307.

Haviland, S. E., &Clark, H. H. (1974). What' new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*, 512–521.

Hintzman, D. L. (1986). "Schema-abstraction" in a multiple trace model. *Psychological Review, 93*, 411–428.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kaschak, M. P, Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., &Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition, 94*, B79–B89

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29*, 133–159.

Kosslyn S. M. (1994). *Image and brain: the resolution of the imagery debate*. Cambridge, MA: MIT Press.

Lakoff, G., & Nunez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York: Basic Books.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211–240.

Langston, W., Kramer, D. C., & Glenberg, A. M. (1998). The representation of space in mental models derived from text. *Memory and Cognition, 26*, 247–262.

Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin and Review, 15*, 838–844.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics, 20*, 193–232.

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology, 15*, 533–545.

McKoon, G., & Ratcliff, R. (1992). Inferences during reading. *Psychological Review, 99*, 440–466.

McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theories, interventions, and technologies*. Mahwah, NJ: Erlbaum.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283–312.

Meteyard, L., Bahrami, B., &Vigliocco, G. (2007) Motion detection and motion words: Language affects low-level visual perception. *Psychological Science, 18*, 1007–1013.

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28*, 293–312.

Murphy, G. L., & Shapiro, A. M. (1994). Forgetting of verbatim information in discourse. *Memory and Cognition, 22*, 85–94.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes, 26*, 131–157.

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 777–784.

Radvansky, G. A., &Copeland, D. E. (2000). Functionality and spatial relations in situation models. *Memory and Cognition, 28*, 987–992.

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11*, 289–312.

Rinck, M., & Bower, G. H. (1995). Anaphora resolution and the focus of attention in situation models. *Journal of Memory and Language, 34*, 110–131.

Robertson, D. A., Gernsbacer, M. A., Guidotti, S. J., Robertson, R. R. W., Irwin, W., Mock, B. J., & Campana, M. E. (2000). Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science, 11*, 255–260.

Roland, D., Dick, F., & Elman, J.L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57*, 348–379.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996) Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories activates neural representations of perceptual and motor experiences. *Psychological Science, 20*, 989–999.

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language, 53*, 125–140.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory and Cognition, 34*, 78–89.

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes, 16*, 3–34.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Zacks, J., Speer, N., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138*, 307–327.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 920–933

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1196–1207.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 35–62). New York: Academic Press.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*, 292–297.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.

Zwaan, R. A., &Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*, 1–11.