Constructing Inferences and Relations during Text Comprehension

Arthur C. Graesser, Peter Wiemer-Hastings, and Katja Wiemer-Hastings

The University of Memphis

Revision: 1/15/99

Send correspondence to:

Art Graesser
Department of Psychology
Campus Box 526400
The University of Memphis
Memphis, TN 38152-6400
(901) 678-2742
(901) 678-2579 (fax)
a-graesser@memphis.edu

TEXT REPRESENTATION: LINGUISTIC AND PSYCHOLINGUISTIC ASPECTS, Sanders, Schilperoord, Spooren (eds.), Amsterdam/Philadelphia: Benjamins, 2001.

Comprehenders make inferences when they read text, watch film, and observe the real world. This intuition is shared by virtually everyone in linguistics (including psycholinguistics, computational linguistics, sociolinguistics, and text linguistics) and in cognitive science (discourse psychology, artificial intelligence, philosophy of mind, social cognition, cognitive anthropology). However, there has been considerable controversy over matters of inference generation. What inferences are generated? When are they generated? What sources of information need to be intact when inferences are generated? What cognitive processes and representations produce inferences during comprehension? And what, precisely, is an inference anyway?

Twenty years ago there was very little scientific knowledge about inferences in text comprehension. Most research efforts concentrated on the representation of explicit text and the process of linking anaphoric expressions (e.g., noun-phrases, pronouns) to previous explicit text constituents. Times have changed in the world of discourse psychology. There have been serious efforts by discourse psychologists to dig deeper and understand how readers construct 'situation models', i.e., mental models of what the text is about. For example, the situation model for a story would consist of a microworld with characters who perform actions in pursuit of goals, events that present obstacles to goals, conflicts between characters, emotional reactions, the spatial setting, the style and procedure of actions, objects, properties of objects, traits of characters, and mental states of characters. Inference generation is inextricably bound to the process of constructing a situation model. The research efforts in recent years have produced a wealth of theoretical positions in discourse psychology, each of which makes distinctive claims about situation model construction and inference generation: The constructionist theory (Graesser, Singer, and Trabasso 1994), the construction-integration model (Kintsch 1998), the structure building framework (Gernsbacher 1997, see chapter by Gernsbacher in this volume), the event indexing model (Zwaan, Langston, and Graesser 1995; Zwaan and Radvansky 1998), the resonance model (Myers & O'Brien, 1998; O'Brien, Raney, Albrecht, and Rayner 1997), the landscape model (van den Broek, Young, Tzeng, and Linderholm 1998), the schema copy plus tag model (Graesser, Kassler, Kreuz, and McLain-Allen 1998), the 3CAPS model (Goldman, Varma, and Cote 1996), and the minimalist hypothesis (McKoon and Ratcliff 1992). Our scientific knowledge of inference generation and situation model construction has evolved from barren to overwhelming in just two decades.

It is beyond the scope of this chapter to introduce and clarify all of the controversies associated with inference research. Our objective is more modest. We want to show how world knowledge plays a central role in the mechanisms that construct inferences and the relations that bind text constituents. Any model will be a dismal failure if it confines its analysis to linguistics and the surface cues in the explicit text. Our chapter is divided into three sections. First, we describe Graesser's constructionist theory of inference generation (Graesser et al. 1994). This theory offers discriminating predictions about what knowledge-based inferences are generated when readers construct a situational model for a text. Second, we briefly describe a three-pronged method for investigating inference generation. This method is needed to provide a rigorous scientific investigation of inference generation. Third, we present a catalogue of relations that are used to connect text constituents and the conceptual entities in world knowledge structures. The catalogue of relations is based on analyses of world knowledge, but the constraints of world knowledge are to some extent manifested in language and rhetoric. The underlying assumption is that world knowledge is sufficiently constrained and systematic that it can be incisively integrated with theories of language processing that have traditionally been confined to the lexicon, syntax, and semantics (see recent book edited by Tomasello 1998).

A Constructionist Theory of Inference Generation

Many inferences are the same when comprehending a sequence of events via text versus film versus observation of the real world. For example, consider the following cryptic sequence of descriptions.

> A diner. A couple sits down at a table. The young woman has a distressed look on her face. She takes a letter out of her purse. She slides it to the young man. The young man lifts up the letter. He reads it. Soon he is stunned. He sits motionless. The tears start to fall. The young woman gets up from the table. She stares at the floor. She leaves.

The 13 verbal descriptions could easily be captured by 13 corresponding film clips without appreciably changing the meaning. There would be appropriate camera angles and distances for capturing the sequence of scenes, actions, and emotional expressions. Similarly, a customer at a nearby table could witness the same sequence of 13 'observations' without appreciably changing the meaning. When comprehending "A couple sits down at the table," comprehenders infer the 'superordinate goals' (motives) of getting food and having a conversation. When comprehending "She slides the letter to the young man", comprehenders infer that the woman has the superordinate goal of getting the man to read the letter. When comprehending "Soon he is stunned", comprehenders infer the 'causal antecedent' state that the letter has disappointing news. This inference is verified further and strengthened when comprehending the subsequent two sentences ("He sits motionless," "Tears start to fall"). Most comprehenders infer the plausible causal antecedent event that the couple breaks up, at or before "She leaves." It would appear, therefore, that the generation of these knowledge-based inferences is governed by mechanisms that transcend the medium; the same inferences occur while reading text, watching film, and observing the real world. In a nutshell, the medium is **not** the message.

We could speculate on the minimum set of statements/clips/observations that would be needed to convey the break-up scenario. The following four might work: The young woman has a distressed look on her face, the young man reads the letter, tears (of the young man) start to fall, and the young woman leaves. The following four would definitely not work: A diner, a couple sits down at the table, the young man lifts up the letter, and the young woman gets up from the table. The minimal set would be the same for text, film, and the world. If we were to rate the statements on importance to the break-up scenario, the ratings would correlate very highly among the text, the film, and the world. So once again, the medium is not the message.

Nevertheless, there are some nontrivial differences between text, film, and the real world. These differences have repercussions on the sort of inferences that would be anticipated in these three modes. A large amount of information is perceptually available to the viewer in the case of films and the real world. The viewer directly observes the spatial layout of the environment, visual features of objects and people, the actions that are performed by people, the manner in which actions are performed, and events that are visually prominent. The viewer does not have to generate inferences about this information that has strong visual support. However, the viewer does need to generate inferences that **explain** the visible actions, events, and states; these inferences include the superordinate goals, causal antecedents, emotions, and perhaps traits of people. These inferences are normally invisible, but they do explain what is seen. The theory of inference generation that we will be advocating, called the 'constructionist theory' (Graesser et al. 1994; Graesser and Zwaan 1995; Graesser and Wiemer-Hastings in press), assumes that these 'explanation-based' inferences are routinely generated while comprehending information in all modes -- text, film, and the world. An adult comprehender has had years of experience (16 hours per day) generating explanation-based inferences, so these strategies of generating inferences are well weathered and automatized in the cognitive system. It is very adaptive for a person to construct these explanation-based inferences because they are associated with the achievement of goals, the causes and consequences of obstacles, and survival in the social and physical world (Mooney 1990; Nordman and Vonk 1998; Schank 1986).

The constructionist theory offers discriminating predictions about the classes of inferences that adults make when they comprehend text. Some classes of inferences are made routinely and quickly, whereas others are made sporadically or are entirely missed unless the reader has an idiosyncratic goal to focus on a particular dimension of knowledge. When adults comprehend text, for example, they routinely make the explanation-based inferences, but they do **not** consistently generate many other categories of inferences, such as the spatial layout, visual features of objects and people, and the style or method of performing actions. The latter information receives strong visual support when films are viewed and the real world is observed, so the inference strategies during text comprehension are minimally developed in these arenas. According to the constructionist theory, adult readers generate the explanation-based inferences consistently and quickly, whereas the construction of the other classes of inferences mentioned above (called 'elaborative' inferences) is much more variable and time-consuming. For example, it would take several minutes and considerable cognitive resources to construct a mental map of a region with people and objects distributed in a particular layout. Therefore, it is unlikely that a detailed mental map is constructed while reading a text at the normal rate of 250-400 words per minute. Yet the mind can perceive a spatial layout in less than a second when viewing a film or the real world.

It is important to acknowledge that the predictions of the constructionist theory are not equivalent to alternative theories of inference generation in discourse psychology. For example, McKoon and Ratcliff’s (1992) minimalist hypothesis does not predict that explanation-based inferences are routinely made. The minimalist hypothesis instead predicts that the only inferences that readers routinely and quickly generate are those that are readily available in working memory and that are needed to establish local text coherence. Early theories of mental models (Glenberg, Meyer, and Lindem 1987; Johnson-Laird 1983) predicted that readers construct spatial inferences, whereas the constructionist theory predicts that it is too time-consuming to construct a rich spatial situation model during normal reading. Graesser et al. (1994) reviews the evidence for the predictions of the constructionist theory and how its predictions are different from alternative theories, models, and hypotheses in discourse psychology. However, it is beyond the scope of this chapter to compare and contrast the various theoretical predictions.

Text, film, and the world differ on another important dimension. Texts and films are intentionally created by humans in an effort to convey messages to the comprehender. The information must be delivered in a coherent fashion that conveys the message (e.g., point, moral, macrostructure). In the case of text, the pragmatic communicators of the message are the writer and the narrator. In the case of film, there is the script writer, the director, and the camera-operator. However, this pragmatic 'message level' is absent in the real world. People in the real world do not enact scripts in the service of messages to viewers. They just live their lives, and much of life is rather uneventful. Thus, the content of texts and films is intentionally and coherently constructed in the service of a message, whereas these constraints are absent in the real world. According to the constructionist theory of inference generation, comprehenders of text attempt to a construct a meaning representation and supporting inferences in a fashion that achieves 'coherence' at local and global levels. Readers generate inferences that fill gaps in the main messages. They generate inferences that explain why the writer bothers to mention something that otherwise would be insignificant (e.g., an 'out of the blue' clue in a mystery novel).

This attempt to construct local and global coherence is appropriate when comprehenders process text, film, and other communication artifacts. Such mechanisms are strategic in the sense that the reader invokes cognitive procedures deliberately, systematically, and continuously. In contrast, a search for coherence is not an appropriate strategy when viewing life's activities because much of life does not unfold in coherent packages. Consequently, adults have much less experience generating coherence-based inferences; such strategies are acquired only when they read text and view films. Coherence-based inferences may not be routinely made during comprehension by readers who rarely read texts and view films because they lack the experience needed to overlearn such comprehension strategies. Adults who read infrequently may have trouble constructing inferences that address the global message of a text and the pragmatic level (e.g., the motives and attitudes of the writer).

The constructionist theory stipulates that readers generate two other classes of inferences. These inferences are difficult to predict ahead of time because they are contingent on the detailed composition of the text and the mental state of the reader. 'Passive-activation' inferences are encoded because they are activated and reactivated by multiple sources of information (e.g., words, propositions, contents of working memory, scripts, global macrostructures), or are strongly activated by one important information source. There are several psychological models that make specific predictions about what inferences are activated and encoded through passive activation mechanisms, such as Kintsch's construction-integration model (1998), the resonance model (Myers and O’Brien 1998; Myers et al. 1994), and the minimalist hypothesis (McKoon and Ratcliff 1992). 'Reader-goal' inferences are motivated by the idiosyncratic goals that the reader has while reading the text. For example, if the reader has the goal of tracking the personality of one of the characters, then there will be many inferences about that character's personality. If the reader wants to trace the spatial layout of a room in a short story, then relevant spatial inferences are generated. There is ample experimental evidence that readers generate spatial inferences if the instructions or the experimental task encourage them to construct a spatial mental model (Glenberg, Meyer, and Lindem 1987; Morrow, Greenspan, and Bower 1987; Rinck, Williams, Bower, and Becker 1996), whereas these spatial inferences are rarely constructed when naturalistic stories are comprehended without goals that monitor spatial processing (Zwaan and Van Oostendorp 1993).

The constructionist theory stipulates that many other classes of inferences are **not** routinely constructed during comprehension. For example, readers do not normally generate the 'logic-based' inferences that are derived from most of the well-formed rules in a syllogistic reasoning task, such as *modus tollens* [i.e., (A implies B) and (not B), therefore (not A)] and DeMorgan's rule. It takes time (measured in minutes) and an external memory (such as truth tables or Venn diagrams) to construct such analytical inferences, so they will not be made at a reading rate of 250-400 words per minute. Humans apparently can handle *modus ponens* [i.e., (A implies B) and A, therefore B] because this rule is easy on the cognitive system; the expressions do not involve negation and the inferences can be produced by pattern-triggered associative operations. Readers do not generate most 'statistics-based' inferences during comprehension because such inferences would take minutes or hours to construct, even by those with high expertise in statistics. Some readers can accommodate simple statistical inferences, such as averages or ranges on a small set of numbers. Readers are not very good at generating 'causal consequence' inferences (or expectations) that forecast what should happen many steps into the future. Readers can sometimes predict what will immediately occur after a text event, particularly if the context is so constrained that there are only one or two likely consequences. However, most expectations end up getting disconfirmed by future occurrences, particularly in a dynamic world.

The picture that we have sketched so far suggests that adult readers routinely generate only a small set of inferences, rather than promiscuously generating many classes of inferences. Readers routinely generate inferences that involve explanations, passive-activations, readers' goals, and coherence (the latter perhaps being confined to readers who have high reading fluency). Readers do not tend to generate inferences that involve detailed elaborations (e.g., spatial layout, visual features of objects and people, the manner in which actions and events occur), distant causal consequences, logical syllogisms, and statistics. Instead of assuming that "anything goes", there are constraints from cognition, the world environment, and perhaps biology that limit the classes of inferences that are generated. The constructionist theory offers one position that specifies what inferences are most likely to be generated and what aren't.

A Three-Pronged Method of Investigating Inference Generation

Inferences are not directly manifested in the text, so there needs to be a method of exposing the inferences and testing whether they are generated during normal reading. A three-pronged method has been advocated for studies of inference generation (Graesser et al. 1994; Magliano and Graesser 1991; Suh and Trabasso 1993; Trabasso and Magliano 1996). The three prongs are (1) theoretical predictions, (2) collection of verbal protocols, and (3) collection of on-line behavioral measures.

Theoretical predictions

This prong simply asserts that it is important to identify what classes of inferences are predicted to be generated by the various theoretical models. For example, the constructionist theory would predict that the explanation-based inferences are routinely generated, but not the detailed elaborative inferences. In contrast, McKoon and Ratcliff's (1992) minimalist hypothesis would predict that neither of these classes of inferences are routinely generated. As mentioned above, there is no lack of theories and models in discourse psychology to test.

Verbal protocols

Verbal protocols are collected from a normative group of readers as they comprehend the text, sentence by sentence. These verbal protocols expose the inferences that surface to consciousness when the reader has the luxury of thoughtful reflection on the text. If an inference appears in the verbal protocols, then the experimenter is on safe grounds in arguing that the readers have a sufficient amount of background knowledge to produce the inference. There is a methodological danger in experimenters generating their own inference test items and presuming that readers have the knowledge to generate such inferences. The most common form of verbal protocol is the 'think aloud' protocol. Readers think aloud and express what ever comes to mind as they comprehend the text, sentence by sentence (Cote, Goldman, and Saul 1998; Trabasso and Magliano 1996; Zwaan and Brown 1996). In question answering tasks, the readers answer particular questions about each sentence, such as why, how, and what-happens-next (Graesser and Clark 1985). In question asking tasks, the readers ask questions that come to mind about each sentence (Olson, Duffy, and Mack 1985). All of these verbal protocol techniques expose the inferences that readers make with minimal time constraints.

Analyses of think aloud protocols have confirmed some of the predictions of the constructionist theory. Most of the content of the think aloud protocols consists of explanations, whereas causal consequences and elaborations are considerably less frequent. This pattern of results has occurred for simple children's stories (Trabasso and Magliano 1996), literary stories for adults (Zwaan and Brown 1996), and expository text (Cote et al. 1998). As predicted by the constructionist theory, the inferences that rise to consciousness should include superordinate goals of characters, causal antecedents, and other forms of explanations, whereas there should be a lower incidence of elaborative associations and distant causal consequences (i.e., expectations, forecasts, predictions).

Verbal protocol analyses are also used to expose the content of the information sources that supply the inferences. In essence, verbal protocols are used to perform knowledge extraction (or knowledge engineering) on the packages of generic world knowledge that supply the inferences in the situation model. Many of the knowledge-based inferences are inherited from the word concepts in the explicit text, particularly the nouns, verbs, and adjectives. According to the estimate of Graesser and Clark (1985), approximately 63% of the knowledge-based inferences in stories are inherited from the word concepts. An additional 9% of the inferences are inherited from global concepts that are related to the story, but are not explicitly stated (such as the concepts of fairytale and conflict). The remaining 28% of the story inferences are 'novel-situational' inferences; the content of these inferences did not match any information units in the word concepts or the global concepts.

Graesser and Clark's analysis of the verbal protocols was analogous to research on the phenomenon of conceptual combination. The basic challenge in this research area is to explain the extent to which the representation of a combined concept (e.g., "lamp oil") can be derived from the constituent word concepts ("lamp", and "oil")(Hampton 1987; Rips 1995; Wisniewski 1997). One straightforward approach to investigating conceptual combination is to have participants list attributes of concept A, concept B, and the combined concept AB. To what extent do the attributes listed in the combined concept AB match the attributes listed in A and B? What sort of attributes in a generic concept (such as "oil") end up appearing in the attributes of diverse concept combinations (such as "corn oil" versus "lamp oil" versus "baby oil")? What are the emergent attributes in combination AB that cannot be inherited from either concept A or concept B? It is beyond the scope of this chapter to discuss the principles and theories of conceptual combination. The point we wish to emphasize is that an analysis of these verbal protocols provides a rich data base for inducing the mechanisms of conceptual combination and for testing models that make distinctive predictions. Moreover, Graesser and Clark (1985) adopted a similar approach by using verbal protocols to explore the inferences that are inherited from generic concepts versus those that are novel (emergent) in the situation model.

On-line behavioral measures

The comprehender has the luxuries of time and thoughtful reflection when verbal protocols are collected. Just because an inference appears in a think aloud protocol does not necessarily mean the inference is encoded during normal comprehension. Readers do not have the benefits of reflection and time while reading at a rate of 250-400 words per minute. Therefore, some of the inferences exposed by the verbal protocols may not be generated on-line during normal reading. Conversely, some of the inferences encoded during normal reading may not appear in the verbal protocols; such inferences may be difficult to express in words or may not be amenable to conscious introspection. Verbal protocols do not provide a perfect window to the inferences that get constructed during normal reading, so it is important to collect on-line behavioral measures in rigorous tests of an inference theory.

Discourse psychologists have explored a large number of measures and tasks that tap on-line comprehension processes and inference generation (Graesser, Millis, and Zwaan 1997; Haberlandt 1994). For example, the typical dependent measures include self-paced reading times for text segments (e.g., words, clauses, sentence), fixation times on words during eye-tracking, lexical decision latencies on test words (i.e., whether a test string is a word or a nonword), naming latencies on test words, latencies to verify whether a test sentence is true or false, latencies to decide whether a test segment has been presented earlier (i.e., recognition memory), and speeded recognition judgments under a deadline procedure. All of these methods have been used in our investigations of inference generation, but most of our studies have investigated sentence reading times (Graesser and Bertus 1998), lexical decision latencies (Long, Golding, and Graesser 1992; Magliano, Baggett, Johnson, and Graesser 1993; Millis and Graesser 1994) and word naming latencies (Long et al. 1992). The results of these studies have been compatible with the results of the think aloud studies and the constructionist theory: the explanation-based inferences are encoded more quickly and strongly during comprehension than are distant causal consequences and frivolous elaborative inferences.

One line of research will be described in order to illustrate the careful experimental control that was imposed on some tests of the constructionist theory. After each sentence in a text was read, a test string appeared (cued with asterisks or in some other fashion). Some of the test strings matched a distinctive noun, verb, or adjective in the inference under consideration. For example, "eat" would be the inference test word from the inference "the couple wanted to eat." This inference would be theoretically predicted while reading the second sentence of the example break-up scenario ("A couple sits down at a table."). Test words were prepared for inferences in different theoretical categories, such as superordinate goals, subordinate actions, causal antecedents, causal consequences, and elaborative states. The samples of these test words were equilibrated on measures derived from verbal protocols that were collected from other groups of participants. For example, if 25% of the participants had expressed the superordinate goal inferences in the verbal protocols, then 25% would also have expressed the subordinate actions. Similarly, the samples of test words were equilibrated on the likelihood that they matched information units from the word concepts that appeared in the explicit sentences. For example, consider the word concepts that appeared in the text up through the second sentence in the break-up scenario: diner, couple, sit, and table. Knowledge extraction methods had been used (from previous groups of participants) to expose the content of each of these generic word concepts. The test inferences in the different classes were equilibrated on the likelihood that they matched content in the four word concepts.

There were additional experimental controls in these studies. We imposed careful control over comprehension time by using a rapid serial visual presentation task. The sequence of words in the sentence are presented for a short duration (250-500 milliseconds per word): A, couple, sits, down, at, a, table. The time-course of inference activation is traced by presenting the inference test word ("eat") at specific durations after the end of the sentence (e.g., 150, 500, 1000, versus 2000 milliseconds). An inference is assumed to be encoded if its test word has a comparatively short latency and this facilitation is sustained over time (rather than quickly decaying). An inference is assumed to be activated if its latency in the inference context is shorter than the same word in the context of another text. Thus, an inference activation score is measured as [latency (unrelated context) - latency (inference context)].

These well controlled experiments supported the predictions of the construction theory. That is, the explanation-based inferences had higher inference activation scores than did the distant causal consequences and the elaborative inferences (Long et al. 1992; Magliano et al. 1993; Millis and Graesser 1994). The results of the experiments that collected on-line behavioral measures were therefore compatible with the verbal protocol studies.

A Catalogue of Relations based on World Knowledge

According to the constructionist theory, readers attempt to achieve local and global coherence when they comprehend text. The coherence is sometimes driven by explicit features of the text, such as anaphoric references, connectives, transitional phrases, rhetorical predicates, and signaling devices. However, sometimes the coherence relations are constructed inferentially. Good readers are able to infer the appropriate coherence relations that bind text constituents. It is important to acknowledge that there is no guarantee that coherent text representations are constructed because the process is contingent on the reader's judgment that the author intended to construct a coherent message. All bets are off if the text is so disjoint and inconsiderate that the reader gives up trying to construct a coherent message. However, most naturalistic texts have some semblance of coherence and most readers do make the effort to achieve a coherent construction.

Coherence relations have been extensively analyzed in the fields of text linguistics and discourse processes (for example, consider the chapters in the present volume by Chanquoy, by Hobbs and Redeker, by Knott, and by Henkemans). Researchers in these fields have proposed taxonomies of coherence relations that allegedly are needed to explain the structure and processing of oral and written discourse (Givon 1993; Halliday and Hasan 1976; Mann and Thompson 1986; Sanders 1997; Sanders, Spooren, and Nordman 1992). For example, a relatively small set of coherence relations appears to underlie the connectives that explicitly occur in texts. There are connectives that signify ‘temporality’ (e.g., before, after, during, and, then), ‘causality’ (because, so), ‘intentionality’ (in order to, for the purpose of), ‘opposition’ (but, however, on the other hand), ‘logical implication’ (therefore, thus), and so on.

A text is regarded as coherent to the extent that its explicit statements can be connected to each other conceptually. Local coherence is achieved if the reader can connect the incoming sentence to information in the previous sentence or to the content in working memory. Global coherence is achieved if the incoming sentence can be connected to the text macrostructure (i.e., major message or point) or to information much earlier in the text that no longer resides in working memory. Available research in discourse psychology suggests that readers attempt to achieve coherence at both local and global levels (Albrecht and O'Brien 1993; Myers, O'Brien, Albrecht, & Mason, 1994; Singer, Graesser, and Trabasso 1994; van den Broek and Lorch 1993), although there is some debate about the consistency in which global coherence is achieved.

The establishment of text coherence is sometimes facilitated when there are explicit connectives that signal how text constituents should be connected (Britton and Gulgoz 1991; Deaton and Gernsbacher in press; Millis and Just 1994). However, explicit textual cues that signal coherence relations are not always necessary for the establishment of conceptual coherence because these links can sometimes be filled in inferentially during the construction of the situation model. The reader requires fewer explicit cues to the extent that there is ample world knowledge about the content of the sentences.

Zwaan has proposed an 'event indexing' model that accounts for the reader's construction of a multithreaded situation model while reading stories (Zwaan, Magliano, and Graesser 1995; Zwaan, Langston, and Graesser 1995; Zwaan and Radvansky 1998). This model assumes that the reader monitors five conceptual dimensions: protagonists, temporality, spatiality, causality, and intentionality. A break in continuity may exist on any of these dimensions when an incoming sentence is read. Protagonist discontinuity exists when an incoming event E has a character that is different from the characters in the previous sentence P. A temporal discontinuity exists when event E occurs much later in time than P, or is part of a flashback. A spatial discontinuity exists when event E occurs in a spatial setting that is different from P. A causal discontinuity exists when event E is not causally related to P. An intentional discontinuity exists when event E is part of a character's plan that is different from any plans associated with P. Research by Zwaan supports the claim that reading times for event E increase as a function of the number of dimensional discontinuities that exist. Also, a discontinuity on any one of these dimensions significantly increases reading time.

World knowledge plays a central role during the process of establishing continuities on these dimensions, as well as other types of coherence relations. Consequently, we believe that it is worthwhile to take stock of the relations that have been identified by researchers who investigate the representation of world knowledge. In addition to the coherence relations that have been identified in text linguistics and discourse processes, the fields of artificial intelligence and cognitive science have devoted 30 years of research investigating the conceptual structures that represent world knowledge (Graesser and Clark 1985; Lehmann 1992; Schank and Reisbeck 1981). These conceptual structures consist of a set of nodes (concepts, states, events, actions, goals) that are connected by relational arcs (is-a, has-as-parts, cause, reason, etc.). The relations that exist in world knowledge structures presumably are very relevant to the coherence relations that connect text constituents.

We recently developed a catalogue of relations in a project funded by the Office of Naval Research. This relation catalogue is presented in Table 1. The objective of the research project was to build a computer tool that guides the elicitation of knowledge from expert tactical planners in the military (Williams, Hultman, & Graesser, 1998). This knowledge is extracted during the process of building expert systems. There is a broad landscape of knowledge in tactical planning: planning networks, causal networks, conflict scenarios, organizational hierarchies, friend-foe networks, taxonomic hierarchies, spatial structures, visual descriptions, and logical structures. Although the project focused on tactical planning, the catalogue of relations was expected to adequately cover a large spectrum of knowledge domains.

If a researcher wanted to be extremely precise, there would be literally thousands of relation categories that capture very subtle semantic distinctions. However, a precise analysis runs the risk of being so subtle that a knowledge engineer or discourse analyst would have trouble remembering and applying it; there would be low agreement between a pair of researchers as to whether a particular relation is relevant. It is widely acknowledged that interjudge agreement decreases when there are a large number of categories and the theoretical distinctions are extremely fine-grained (D'Andrade & Wish, 1985). At the other end of the continuum, there are liabilities to having a very small number of relation categories; the set of categories would be so crude that the judges end up glossing over critical theoretical distinctions. The relation categories in Table 1 approach an optimal point in this tradeoff. It is possible to obtain satisfactory interjudge agreement by researchers trained to use these categories. At the same time, the categories are functionally useful in computational models of question answering, summarization, recall, and planning (Graesser and Clark 1985; Graesser, Gordon, and Brainerd 1992).

The relation catalogue in Table 1 contains 22 basic relations altogether. For each relation, there is a definition, a composition rule, and an example. For some of these basic relations, there are additional relations that are synonyms, inverses, negations, and subtypes. Therefore, this scheme could be expanded to over 100 relations after considering the synonyms, inverses, subtypes, and negations of the 22 basic relations. The catalogue accommodates all of the relations that are included in Graesser, Gordon, and Brainerd’s (1992) analysis of world knowledge and most of relations that were reported by the 35 contributors to Lehmann’s edited volume *Semantic Networks in Artificial Intelligence* (1992). We did not include relations that are involved in sentence syntax and in case structure thematic roles (e.g., agent, object, recipient, etc.). A large subset of the relations in Table 1 are also relations that exist in Wordnet (Miller 1990). Wordnet is a large lexicon of nouns, verbs, and adjectives which contains syntactic and semantic features that reflect the use of words in language.

Each of the relations in Table 1 connect two nodes. Each node is assigned to one of five categories:

Concept (C). A noun-like concept, such as "captain," "ship," and "battle."

State (S). An ongoing characteristic that remains unchanged within a relevant time frame, such as "the ship has missiles" and "the water is salty."

Event (E). A state change that occurs within a relevant time frame, such as "the ship sank" and "the fleet threatened the enemy."

Goal (G). A state or event that an agent wants to achieve, such as "the captain wants the ship to reach the port" and "the captain wants to communicate with the enemy."

Style (Sy). The qualitative manner or intensity in which an event unfolds, such as the ship moved "in a zigzag path" and the missile moved "slowly."

It should be noted that the 'intentional actions' of agents are not primitive node categories in this analysis. Instead, they are amalgamations of goal nodes linked to states, events, or style specifications which signify a positive outcome. Therefore, the action “the captain sank the ship” is an amalgamation of two nodes:

(G: captain wanted to sink the ship) ---OUTCOME---> (E: the ship sank).

There is another constraint that the captain executed some plan that led to the sinking ship. Our analysis segregates goals from the events in the world because it is very important in planning to differentiate plans (which may or may not be implemented) from events that actually occur in the world.

Each relation has a ‘composition rule’ which specifies the node categories that can be linked by a particular type of relation. For example, the IS-A relation can link only concept nodes, whereas the REASON relation can link only goal nodes. "Any" signifies that any of the five node categories may apply. Braces {} signify that a set of node categories may occur. For example, {E | S | Sy} signifies that a node may be either an event, a state, or a style specification.

The position that we are advocating is that the relations in this catalogue play an important role in connecting explicit text constituents and in building situation models during comprehension. The relations in Table 1 are in the structures when coherent meaning representations are built. Regarding inferences, some relations play a more important role than others. According to the constructionist model, readers attempt to explain why events, actions, and states occur, so the important conceptual relations to monitor during inference generation are CAUSE, INITIATES, OUTCOME, REASON, and TEMPORAL. In contrast, relations such as HAS-AS-PART, MANNER, PROPERTY, QUANTITY, SPATIAL, and SUBPROCESS are not as prevalent in the inference mechanisms.

An analysis of world knowledge is needed in order to build a complete psychological theory of how humans build coherent messages. Instead of starting with text and language, and asking what text connections are explicitly articulated, we start with world knowledge and ask what relations are prevalent when we make sense of the world. Instead of specifying what coherence relations are needed to bind text constituents, one can inquire what coherence relations are needed in order to build a computational model that performs a variety of processing tasks (such as question answering, recall, summarization, and planning). Of course, what ultimately is desired is a theory of comprehension that specifies how the meaning representations are constructed on the basis of both world knowledge and the surface linguistic cues. A productive research direction for future research would be to specify detailed mappings among (a) surface linguistic cues, (b) world knowledge structures, and (c) cognitive processes.

The contributors to this volume have explored the cross fertilization of research in linguistics and psychology. This interdisciplinary effort is difficult. There is a strong complacent inclination for the fields to remain isolated and insulated from each other. Nevertheless, the efforts will hopefully yield fresh insights. If our claims in this chapter are on the mark, it will also be worthwhile to add computational linguistics and artificial intelligence to the fold.

Table 1: Catalogue of Relations

(1) **AND**
Definition: Both A and B exist or occur.
Composition rule: {Any} <---AND---> {Any} [Note: The arc is bi-directional.]
Example: (C: ships) <-- AND--> (C: planes)

(2) **BECOMES**
Definition: One concept changes into another concept.
Composition rule: (C) ---BECOMES---> (C)
Example: (C: civilian) ---BECOMES---> (C: enemy)

(3) **CAUSE**
Synonyms: **CONSEQUENCE**
Subtypes: **DIRECT-CAUSE, ENABLES**
Inverse: **PRECONDITION, PREREQUISITE**
Definition: A directly or indirectly causes or enables B. The beginning of A precedes the beginning of B.
Composition rule: { E | S | Sy } ---CAUSE---> { E | S | Sy }
Example: (E: The plane flew too low) ---CAUSE---> (E: Radar flagged the plane)

(4) **CONTROLS**
Synonym: **SUPERVISES**
Definition: A has direct control over the activities of B.
Composition rule: (C) ---CONTROLS---> (C)
Example: (C: captain) ---CONTROLS---> (C: lieutenant)

(5) **DISABLES**
Synonyms: **BLOCKS**
Definition: A stops or prevents B.
Composition rule: {G | E | S | Sy} ---DISABLES---> {G | E | S | Sy}
Example: (S: ship has no fuel) ---DISABLES---> (E: ship moves to island)

(6) **FRIEND (and FOE)**
Synonyms: **LIKES, ALLY**
Negation: **FOE, ENEMY,DISLIKES**
Definition: One animate entity is in an alliance with another animate entity.
Composition rule: (C) <---FRIEND---> (C) [Note: The arc is bi-directional.]
Example: (C: Iran) <---FOE---> (C: Iraq)

(7) **HAS-AS-PART**
Synonyms: **HAS-COMPONENT**
Inverses: **IS-A-PART-OF**
Definition: A has a part or component B.
Composition rule: (C) ---HAS-AS-PART---> (C)
Example: (C: The U.S.S. Nimitz) ---HAS-AS-PART---> (C: flight deck)

(8) **IMPLIES**
Synonyms: **IF-THEN**
Definition: If A exists or occurs, then B exists or occurs. A and B overlap in time.
Composition rule: {Any} ---IMPLIES---> {Any}
Example: (S: The fleet has many ships) ---IMPLIES---> (S: The fleet is powerful)

(9) INITIATES

Synonyms: **ELICITS**
Inverses: **CONDITION, CIRCUMSTANCE, SITUATION**
Negation: **DISABLES**
Definition: A initiates or elicits a goal.
Composition rule: { E | S | Sy } ---INITIATES---> (G)
Example: (S: The ship was low on fuel) ---INITIATES---> (G: captain get fuel for ship)

(10) IS-A

Subtypes: **IS-A-KIND-OF, IS-AN-INSTANCE-OF, IS-A-MEMBER-OF, IS-A-TYPE-OF**
Synonyms: **SUPERCONCEPT**
Inverses: **KIND, INSTANCE, MEMBER, SUBTYPE, SUBCONCEPT**
Negation: **IS-NOT-A, OPPOSITE**
Definition: A is a subcategory or instance of B.
Composition rule: (C) ---IS-A---> (C)
Example: (C: The U.S.S. Nimitz) ---IS-A---> (C: ship)

(11) MANNER

Definition: A specifies the manner in which a state change B occurs or a goal is achieved. A and B overlap in time.
Composition rule: { E | Sy } ---MANNER---> {E | Sy}
(G) ---MANNER---> (G)
Example: (E: the troops moved) ---MANNER---> (Sy: quickly)

(12) NAME

Definition: A concept or other node is named with a particular label.
Composition rule: {Any} ---NAME---> (name label)
Example: (E: USA attacks Iraq) ---NAME---> ("Desert Storm")

(13) OR

Definition: Either A or B exist or occur, but not both
Composition rule: {Any} <---OR---> {Any} [Note: The arc is bi-directional.]
Example: (G: fleet threaten enemy ship) <---OR---> (G: fleet destroy enemy ship)

(14) OUTCOME

Synonyms: **RESULT**
Definition: A specifies whether or not the goal B is achieved
Composition rule: (G) ---OUTCOME--- > { E | S | Sy }
Example: (G: commander destroy ship) ---OUTCOME---> (E: ship was destroyed)

(15) PROPERTY

Synonyms: **ATTRIBUTE, CHARACTERISTIC, FEATURE**
Definition: A concept has a particular characteristic.
Composition rule: (C) ---PROPERTY---> { E | S | Sy }
Example: (C: ship) ---PROPERTY---> (S: the ship is buoyant)

(16) QUANTITY

Synonyms: **NUMBER, FREQUENCY**
Definition: The number of instances of a concept or event.
Composition rule: {C | E}---QUANTITY---> (number)
Example: (E: The gun fired) ---QUANTITY---> (12 times)

(17) REASON

Synonyms: **PURPOSE, MOTIVE**
Inverses: **METHOD, PLAN, STEP, PRE-ACTION**
Definition: Goal A is the reason or motive for implementing a method, plan, or action B. The outcome of B is achieved before the outcome of A. A is a superordinate goal of B.
Composition rule: (G) ---REASON---> (G)
Example: (G: commander fire missile) ---REASON---> (G: commander destroy ship)

(18) REFERENTIAL POINTER (RP)

Synonyms: **REFERS-TO**
Definition: One node refers to another set of nodes.
Composition rule: {Any} ---RP---> { Any}
Example: (C: battle) ---RP---> [(E: ship 1 fires at ship 2), (E: ship 2 fires at ship 1)]

(19) SIMILAR-TO

Subtypes: **EQUIVALENT**
Negation: **DISSIMILAR, CONTRASTS,CONTRADICTS**
Definition: Two nodes are very similar in content or features.
Composition rule: {Any} ---SIMILAR-TO----> {Any}
Example: (C: ship 1) ---SIMILAR-TO---> (C: ship 2)

(20) SPATIAL RELATION

Definition: One spatial region, object, or part is spatially related to another.
Composition rule: (C) ---SR---> (C)
Example: (C: mast) --ABOVE---> (C: deck)
The following subtypes of spatial relationships are self-explanatory:

ABOVE	**BELOW** (inverse)
BESIDE	**BY** (synonym)
BETWEEN	
CONTAINS	**IS-IN** (inverse)
CONNECTED-TO	
EAST-OF	**WEST-OF** (inverse)
INSIDE-OF	**OUTSIDE-OF** (inverse), **WITHIN** (synonym)
LEFT-OF	**RIGHT-OF** (inverse)
NEAR	
NORTH-OF	**SOUTH -OF** (inverse)
ON-TOP-OF	**UNDERNEATH** (inverse), **SUPPORTS** (inverse)
SURROUNDS	**ENCAPSULATES** (synonym)
TOUCHES	**ABUTS** (synonym)

The following relations specify a quantity in 3-dimensional space:

POSITION	C-rule: (C) ---POSITION---> (x,y,z)
ORIENTATION	C-rule: (C) ---ORIENTATION---> (rotated 45-degrees on x,y)

(21) SUBPROCESS

Definition: Event A is a subprocess of event B.
Composition rule: (E) ---SUBPROCESS---> (E)
Example: (E: ship fires missile) ---SUBPROCESS---> (E: chamber releases missile)

(22) TEMPORAL RELATION

Definition: Point or duration A is related to point or duration B in time.

Composition rule: {G | E} ---TR---> {G | E}
Example: (E: missile hits ship) ---BEFORE---> (E: ship sinks)
BEFORE **AFTER** (inverse)
DURING

Author Note

Correspondence concerning this article should be addressed to Arthur C. Graesser, Department of Psychology, Campus Box 526400, The University of Memphis, Memphis, TN 38152-6400, a-graesser@memphis.edu. This research was partially funded by grants to the first author by the Office of Naval Research (N00014-95-11113 and N00014-98-1-0331) and the National Science Foundation (SBR 9720314).

References

Albrecht, J.E., & O'Brien, E.J. (1993). Updating a mental model: Maintaining both local and global coherence. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 1061-1070.

Britton, B.K., & 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, 329-404.

Cote, N., Goldman, S.R., & Saul, E.U. (1998). Students making sense of informational text: Relations between processing and representations. Discourse Processes, 25, 1-54.

D'Andrade, R.G., & Wish, M. (1985). Speech act theory in quantitative research on interpersonal behavior. Discourse Processes, 8, 229-259.

Deaton, J.A., & Gernsbacher, M.A. (in press). Causal conjunctions and implicit causality cue mapping in sentence comprehension. Journal of Memory and Language.

Gernsbacher, M.A. (1997). Two decades of structure building. Discourse Processes, 23, 265-304.

Givon, T. (1993). Coherence in text, coherence in mind. Pragmatics & Cognition, 1, 171-227.

Glenberg, A.M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. Journal of Memory and Language, 26, 69-83.

Goldman, S.R., Varma, S., & Cote, N. (1996). Extending capacity-constrained construction-integration: Toward "smarter" and flexible models of text comprehension. In B.K. Britton and A.C. Graesser (Eds.), Models of understanding text (pp. 73-114). Mahwah, NJ: Erlbaum.

Graesser, A.C., & Bertus, E. (1998). The construction of causal inferences while reading expository texts on science and technology. Journal of the Scientific Studies of Reading, 2, 247-269.

Graesser, A.C., Gordon, S., & Brainerd, L.E. (1992). QUEST: A model of question answering. Computers & Mathematics with Applications, 23, 733-1992.

Graesser, A.C., & Clark, L.C. (1985). Structures and procedures of implicit knowledge. Norwood, NJ:Ablex.

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., Millis, K. K., & Zwaan. R. A. (1997). Discourse comprehension. Annual

Review of Psychology, 48, 163-89.

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

Graesser, A.C., & Wiemer-Hastings, K. (in press). Situation models and concepts in story comprehension. In Goldman, S.R., Graesser, A.C., and van den Broek, P. (Eds). Festschrift for Tom Trabasso. Mahwah, NJ: Erlbaum.

Graesser, A. C., & Zwaan, R. A. (1995). Inference generation and the construction of situation models. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), Discourse comprehension: Strategies and processing revisited (pp. 117-139). Hillsdale, NJ: Erlbaum.

Haberlandt, K. (1994). Methods in reading research. In M.A. Gernsbacher (Ed.), Handbook of psycholinguistics (pp. 1-31). New York: Academic Press.

Halliday, M.A.K., & Hasan, R. (1976). Cohesion in English. London: Longmans.

Hampton, J.A. (1987). Inheritance of attributes in natural concept conjunctions. Memory & Cognition, 15, 55-71.

Johnson-Laird, P.N. (1983). Mental models. Cambridge, MA: Harvard University Press.

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

Lehmann, F. (1992)(Ed.). Semantic networks in artificial intelligence. Oxford, England: Pergamon Press.

Long, D. L., Golding, J. M. & Graesser, A. C. (1992). Test of the on-line status of goal-related inferences. Journal of Memory and Language, 31, 634-647.

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., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. Discourse Processes, 16, 35-53.

Mann,W.C., & Thompson, S.A.(1986). Relational propositions in discourse. Discourse Processes, 9, 57-90.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. Psychological Review, 99, 440-466.

Miller, G.A.(1990). WordNet: An on-line lexical database. International Journal of Lexicography, 3.

Millis, K., & Graesser, A.C. (1994). The time-course of constructing knowledge-based inferences for scientific texts. Journal of Memory and Language, 33, 583-599.

Millis, K.K., & Just, M.A. (1994). The influence of connectives on sentence comprehension. Journal of Memory and Language, 33, 128-147.

Mooney, R.J. (1990). A general explanation-based learning mechanism and its application to narrative understanding. San Mateo, CA: Morgan Kaufman.

Morrow, D.G., Greenspan, S.L., & Bower, G.H. (1987). Accessibility and situation models in narrative comprehension. Journal of Memory and Language, 26, 165-87.

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

Myers, J.L., O'Brien, E.J., Albrecht, J.E., & Mason, R.A. (1994). Maintaining global coherence during reading. Journal of Experimental Psychology: Human Learning, Memory, and Cognition, 20, 876-86.

Noordman, L.G.M., & Vonk, W. (1998). Memory-based processing in understanding causal information. Discourse Processes, 26, 191-212.

O'Brien, E.J., Raney, G.E., Albrecht, J.E., & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. Discourse Processes, 23, 1-24.

Olson, G.M., Duffy, S.A., & Mack, R.L. (1985). Question asking as a component of text comprehension. In A.C. Graesser and J.B. Black (Eds.), The psychology of questions (pp. 219-226). Hillsadale, NJ: Erlbaum.

Rinck, M., Williams, P., Bower, G.H., & Becker, E.S. (1996). Spatial situation models and narrative understanding: Some generalizations and extensions. Discourse Processes, 21, 23-56.

Rips, L.J. (1995). The current status of research on conceptual combination. Mind and Language, 10, 72-104.

Sanders, T.J.M., Spooren, W.P.M., & Noordman, L.G.M. (1992). Toward a taxonomy of coherence relations. Discourse Processes, 15, 1-36.

Sanders, T.J.M. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. Discourse Processes, 24, 119-148.

Schank, R.C. (1986). Explanation patterns: Understanding mechanically and creatively. Hillsdale, NJ: Erlbaum.

Schank, R.C. & Reisbeck, C.K. (1981). Inside computer understanding. Hillsdale, NJ: Erlbaum.

Singer, M., Graesser, A.C., & Trabasso, T. (1994). Minimal or global inference during reading. Journal of Memory and Language, 33, 421-441.

Suh, S.Y., & Trabasso, T.. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. Journal of Memory and Language, 32, 279-300.

Tomasello, M. (1998)(Ed.). The new psychology of language: Cognitive and functional approaches to language structure. Mahwah, NJ: Erlbaum.

Trabasso, T., & Magliano, J.P. (1996). Conscious understanding during comprehension. Discourse Processes, 21, 255-287.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1998). The landscape model of *reading inferences and the on-line construction of memory representations. In H. van Oostendorp and S.R. Goldman (Eds.), The construction of mental representations during reading. Mahwah, NJ: Erlbaum.

Williams, K.E., Hultman, E., & Graesser, A.C. (1998). CAT: A tool for eliciting knowledge on how to perform procedures. Behavior Research Methods, Instruments, & Computers, 30, 565-572.

Wisniewski, E.J. (1997). When concepts combine. Psychonomic Bulletin and Review, 4, 167-183.

Zwaan, R.A., & Brown, C.M. (1996). The influence of language proficiency and comprehension skill on situation model construction. Discourse Processes, 21, 289-327.

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., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 386-397.

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

Zwaan, R.A., & Van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension? Discourse Processes, 16, 125-143.