Discourse and Cognition

Arthur C. Graesser

University of Memphis

and Keith Millis

Northern Illinois University

Send correspondence to:

Art Graesser
Institute for Intelligent Systems
365 Innovation Drive
University of Memphis
Memphis, TN, 38152
901-678-4857
901-678-2579 (fax)
a-graesser@memphis.edu

Graesser, A. C., & Millis (2007). Discourse and cognition. In T. Van Dijk (Ed.), *Discourse studies*.

Discourse comprehension is a very rich, multilevel cognitive activity. Consider what it takes to read a simple news story. The words contain letters, pronunciation patterns, meanings, and sometimes emotional nuances. Sentences have syntactic composition, semantic ideas, and stylistic features. Deep comprehension requires the construction of referents of nouns, a discourse focus, presuppositions, and plausible inferences. The reader needs to distinguish old (given) versus new information in the discourse and to acknowledge implicitly what is shared knowledge among individuals in the discourse community (called the common ground). At more global levels, the reader needs to identify the category of the discourse (e.g., objective reports versus editorials), the rhetorical structure, perspectives of different people, and sometimes the attitude of the writer. The processing of all of these levels is effortlessly achieved at a rate of 150 to 400 words per minute by a proficient adult reader. It is simply astonishing that we can understand anything that we read given the many levels of language and discourse that must be mastered.

The various levels of oral conversation and printed text have been identified by discourse researchers over the years. Discourse researchers in the cognitive tradition typically include five levels (Graesser, Millis, & Zwaan, 1997; Van Dijk, 2008; Van Dijk & Kintsch, 1983):

(1) ***Surface code***. The words and syntax that are expressed verbatim.

(2) ***Textbase***. The semantic representation of the explicit text in a stripped down form that removes surface details with few if any semantic consequences. The core idea units are often called propositions.

(3) ***Situation model***. The referential content of what the discourse is about. It includes explicit and inferred people, objects, states, events, actions, processes, goals, and other types of content, as will be illustrated later.

(4) ***Genre and rhetorical structure***. The type of discourse and its structural composition. Example genres at a course-grain level are description, narration, exposition, and persuasion; a more fine-grained differentiation of narration would distinguish folktales, romance novels, historical fiction, and many other subtypes. Each genre has its distinctive structural components, as in the case of setting+plot+moral in a folktale or claim+evidence in an argument.

(5) ***Pragmatic communication***. The point or message that the speaker/writer wishes to convey to the listener/reader in a conversational turn, coherent message, or longer stretch of discourse in a specific social situation. This includes speech participants' intentions, attitudes, emotions, gestures, and joint action in addition to the explicit verbal expressions. The situated context includes the setting (time and place), the participants (communication roles, social status, identities, social relations), the actions being performed, and the mutual cognitions of participants.

Cognitive researchers have spent over 40 years investigating these levels of representation and how they are constructed in the minds of both children and adults. They do this by adopting rigorous scientific methods that explore cognitive mechanisms during and after comprehension or production. The typical scientific experiments: (a) manipulate the discourse, context, and instructions to the participant, (b) collect various cognitive and behavioral measures (e.g., reading times, eye tracking, think-aloud protocols, word naming, ratings, recall, summarization, question answering), and (c) measure or stimulate brain activities. The field of *discourse processes* analyzes the multiple levels of discourse and investigates how the mind comprehends or produces the discourse by using the scientific method.

Inferences

It is important to emphasize that inferences are needed in order to construct most if not all of the five levels of discourse. Inferences are ideas that the reader generates and attaches to one of the five levels (typically levels 2 through 5). Cognitive researchers have conducted experiments to determine which of these inferences are generated during discourse comprehension. For illustration, consider the following excerpt from the beginning of Ian Rankin's *A Good Hanging*.

> It was the perfect Murder.
>
> Perfect, that is so far as the Lothian and Borders Police were concerned. The murderer had telephoned in to confess, had then panicked and attempted to flee, only to be caught leaving the scene of the crime. End of story. (p. 1)

The following four inferences would potentially be made at particular points in the text, and be part of level 3 (situation model)

(A) *Superordinate goal*: "The murderer wanted to turn himself into the police" is an inference when reading the clause *The murderer had telephoned in to confess*.

(B) *Subordinate goal/action*: "The murderer dialed a telephone number" is an inference when reading the clause *The murderer had telephoned in.*

(C) *Causal antecedent*: "The murderer changed his mind" is an inference when reading the clause *he then panicked*.

(D) *Causal consequence*: “The murderer tried to escape” is an inference when reading the clause *he then panicked*.

Of course, other inferences could probably be generated. For example, you probably inferred that you were reading a novel (perhaps a murder mystery) rather than a newspaper article, although this information was not mentioned. These aspects would be part of level 4 (genre).

Cognitive researchers have investigated the classes of inferences that are routinely made during comprehension, versus those classes that are not typically made. According to the constructionist theory of Graesser, Singer, and Trabasso (1994), for example, adult readers would normally generate inferences A and C, but not B and D. Inferences A and C are answers to why-questions, which are foundational questions that guide discourse levels 3, 4, and 5. So when asked “Why did the murder telephone in?”, a plausible answer would be “in order to turn himself in to the police.” When asked “Why did the murderer panic?”, a plausible answer would be “because he changed his mind and did not want to turn himself in to the police.” Inference B is an answer to a how-question (“How did the murderer telephone in?”), not a why-question. Therefore it would be mere ornamentation and not generated unless the reader had the idiosyncratic goal to monitor characters’ precise actions. Inference D is an answer to a what-happened-next-question (“What happened after the murderer panicked?”), not a why-question, and would not theoretically be routinely generated.

These predictions of the constructionist theory are not universally accepted by cognitive researchers. For example, an embodied cognition framework would predict that subordinate actions (D) are constructed. According to the embodied cognition framework, the reader mentally simulates (i.e., generates, constructs) the actions of characters in the story world in rich detail (Glenberg, & Kaschak, 2002). Resonance and minimalist models would predict that most of these inferences are not constructed unless there are surface and text-based features that accurately trigger the correct memories at the right time (Lea et al., 2008; O'Brien et al., 1998). Discourse psychologists continue to hold heated debates and conduct experiments on what inferences are generated during normal comprehension.

Breakdowns and Misalignments in Multilevel Discourse

Comprehension is not always effortless, fast, and satisfactory. When beginning readers struggle over individual words, reading slows down and deeper levels of comprehension are compromised. Even the most proficient adult reader will struggle with a technical expository text on an arcane topic, such as the ingredients of a prescription drug, a legal document, or installation instructions for a piece of computer software. A successful reader is forced to enact deliberate, conscious, effortful, time-consuming strategies in order to repair or circumvent a reading component that is problematic.

Comprehension can break down at any of the five levels according to a multilevel framework (Pickering & Garrod, 2004). The cause of a breakdown may be attributed to either deficits in the reader (i.e., lack of knowledge or processing skill) or the discourse (e.g., incoherent text, unintelligible speech). Breakdowns can range from a temporary and minor irregularity that captures the comprehender's attention to a full meltdown of comprehension. A breakdown at one level might negatively affect other levels. In addition, the comprehender may attempt to compensate for a breakdown at one level by using information from other discourse levels, world knowledge, or external sources (such as other people or technologies). For example, consider the scenarios below that illustrate breakdowns or minor glitches at various discourse levels, along with the resulting consequences.

(Scenario 1) An immigrant arrives in the United States and does not understand the English language at all. A complete breakdown at discourse level 1 also blocks the deeper levels of 2-5.

(Scenario 2) A child has trouble reading aloud a book on dinosaurs because she stumbles on all of the rare words. She has no trouble reading *Princess and the Pea* because she has heard the story dozens of times. Memory and knowledge compensate for word deficits at level 1.

(Scenario 3) Two homebuyers read a legal document that has lengthy sentences with embedded clauses and many logical operators (*and, or, not, if*). They have only a vague idea what the document explicitly states because of complex syntax and a dense textbase (a deficit at levels 1 and 2). However, the couple signs the contract because they understand the purpose of the document and trust the real estate agent. In this case, levels 4 and 5 circumvent the need to understand levels 1-3 completely.

(Scenario 4) A father and son read the directions to assemble a new desk. They get into an argument on how to connect a handle to a drawer because they don't know which screw to attach to the handle. They have no problem understanding the words and textbase in the directions (levels 1 and 2). They also have no problem understanding the genre and purpose of the document (levels 4 and 5). However, they are confused about the screws in the situation model level (level 3). After a series of arguments, mental simulations, and trial and error problem solving, they manage to understand the assembly directions.

(Scenario 5) A novelist loved reading French novels, but painfully recalls receiving a C in French literature when he was in college. Years later, while reading a book on French history, he realized in a flash what had been missing. He had a great memory and appreciation for the settings and plots of those French novels but never understood the novel's deep meaning because he had missed intentions and attitudes of the authors. French history was needed to reconstruct the authors' intent. In this case, discourse levels 1-4 are intact, but the novelist did not make it to level 5 while taking the course in college.

Constraints and Processing Streams

The above scenarios illustrate how deficits in one or more discourse levels can have profound consequences on comprehension and affiliated cognitive processes, such as attention, consciousness, memory, problem solving, and social interactions. Available cognitive research supports some generalizations about the processing order, interaction, coordination, compensatory mechanisms, and constraints of the different levels of discourse comprehension. Here are some of these generalizations:

(a) **There are soft, statistical constraints rather than hard, brittle rules**. The information that accrues at the different levels gets built and influences other levels probabilistically rather than according to hard and fast rules. According to the construction-integration model developed by Kintsch (1998) there is a construction phase of activating code, memories, and rules to varying degrees, followed by an integration phase that probabilistically converges on a final representation through mutual constraint satisfaction.

(b) **There are bottom-up dependencies of meaning**. The ordering on depth is assumed to be 1 → 2 → 3 → 4 → 5. That is, the lower levels constrain the higher levels more than vice versa. However, the ordering is a tendency rather than being rigid because comprehension is a combination of bottom-up and top-down mechanisms. For example, the genre and rhetorical structure might help guide the construction of the textbase and situation model. Nevertheless, a partial analysis of levels 1-3 are needed to adequately construct the rhetorical structure.

(c) **Breakdowns at one level propagate problems to deeper levels**. A breakdown at one level (L) leaves the previous levels intact but threatens the building of level L and the higher levels. As illustrated in scenario 1, if the reader cannot construct a textbase at level 2, there will be problems constructing an adequate situation model, genre, and pragmatic communication, even though level 1 is intact. The individual in scenario 5 made it through level 3, but had problems at level 4 and never achieved level 5 while in college. Comprehenders normally achieve the deepest level of comprehension that is supported by the discourse, their knowledge, and their processing skill, but may get blocked at any level in the processing stream.

(d) **Novel information requires more processing effort than familiar and automatized components**. Novelty of information is a foundational cognitive dimension that attracts attention and effort and that is salient in memory. It could be argued that rare words, the textbase level, and the situation model tend to have the highest density of novel information. In contrast, most aspects of levels 1, 4, and 5 tend to have components that are frequently experienced and therefore well learned and automatized. Reading time studies confirm that more processing time is allocated to rare words than high frequency words (Just & Carpenter, 1987) as well as to new information expressed in the textbase and situation model to a greater extent than old information already mentioned (Haberlandt & Graesser, 1985; Millis, King, & Kim, 2001). In contrast, levels of genre, rhetorical structure, and author characteristics are normally familiar structures that are invisible to the comprehender unless there are irregularities or breakdowns.

(e) **Attention, consciousness, and effort gravitate toward breakdowns and misalignment among levels**. Breakdowns at any level of analysis are likely to draw cognitive resources. Reading time studies have shown that that extra processing time is allocated to pronouns that have unresolved or ambiguous referents (Rayner, 1998), to sentences that have breaks in cohesion (Gernsbacher, 1990), to sentences that have coherence breaks in the situation model (Zwaan & Radvansky, 1998), and to sentences that contradict ideas already established in the evolving situation model (O'Brien, Rizzella, Albrecht, & Halleran, 1998). Attention drifts toward sources of cognitive disequilibrium, such as obstacles, anomalies, discrepancies, contradictions, and cohesion gaps (Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005).

(f) **Breakdowns may be repaired or circumvented by world knowledge, information at other discourse levels, or external sources**. The scenarios illustrate some compensatory mechanisms that repair or circumvent the misfires. World knowledge fills in the lexical deficits in scenario 2, whereas the syntax and textbase deficits in scenario 3 are circumvented by the information in discourse levels 4 and 5. The gaps and misalignments in the situation model of scenario 4 are rectified by extended conversations between father and son and by active problem solving.

The multilevel framework discussed so far provides a plausible sketch of the complexities of constructing meaning at different levels during discourse comprehension. In summary, there are multiple levels of meaning that mutually, but asymmetrically, constrain each other. The components at each level are successfully built if the text is well composed and the reader has prerequisite background knowledge and processing skills. However, there are periodic problems that range from minor misalignments and comprehension difficulties to comprehension breakdowns. These irregularities are magnets of attention that sometimes trigger compensatory mechanisms that repair or circumvent the problems.

Discourse Cognition Embraces Corpus Analysis and Computational Linguistics

Discourse is ubiquitous in the daily activities of members of a culture so there are millions of minutes to learn and to hone the relevant discourse skills. A word, syntactic structure, genre or other discourse component is assumed to be automatized and quickly processed after extensive experience. In contrast, the more novel components are time-consuming, as discussed in the previous section. In order to get a more precise handle on discourse cognition, researchers in the cognitive area have recognized that it is essential for the researcher to get some record of the language and discourse that the community frequently experiences. This need is filled by researchers in the field of *corpus linguistics*. Corpus linguists systematically analyze a large body of naturalistic texts or spoken discourse (called a *corpus*) on various dimensions of language and discourse. The discourse corpus is analyzed by counting the frequency and distribution of discourse elements, categories, features, sequences, global patterns, or combinations of these linguistic/discourse entities. Advances in corpus linguistics are now mainstream in cognitive science and discourse processing.

The corpora analyzed by researchers in discourse cognition are normally quite large, so computers are needed to systematically analyze the printed texts or transcribed oral conversations. During the last decade there have been some revolutionary advances in automating the interpretation of language, discourse, and text on computers. *Computational linguistics* is a hybrid field that integrates computer science, theoretical linguistics, and statistical analyses on discourse corpora. Computers are now able to extract information from texts with a degree of accuracy that convinces some people that computers really can understand the meaning of language, at least to some degree and on some dimensions. For example, an automated essay grader has been developed that can grade essays as reliably as experts in English composition (Landauer, Laham, & Foltz, 2007). There is a system that can classify texts into different discourse genres on the basis of 67 different linguistic and discourse features (Biber 1988). Some of these systems have direct relevance to medicine, mental health, and quality of life. For example, Pennebaker, Booth, and Francis (2007) developed a Linguistic Inquiry and Word Count (LIWC) system that automatically analyzes the words in narratives that are written by victims of traumatic events. Interestingly, the LIWC system can predict how well the person can cope with the trauma and the number of visits the person took to a medical doctor. The Question Understanding Aid (QUAID, www.psyc.memphis.edu/quaid.html) identifies problematic survey questions that have difficult words, complex syntax, or overload working memory (Graesser, Cai, Louwerse,

& Daniels, 2006). The questions that QUAID identifies as problematic have been shown to influence patterns of eye movements and to lower the reliability of humans answering questions on surveys. It should be noted that psychologists or discourse researchers developed all of these computer systems that automatically analyze discourse.

One computer tool called Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004, http://cohmetrix.memphis.edu) analyzes language and discourse on levels 1 through 4. There are over 60 measures of discourse on the public website, whereas the internal research version has nearly a thousand measures at various stages of testing. The Coh-Metrix tool is easy to use and is free to the public. The user simply enters text in a window, clicks a button, and then the system produces dozens of measures of the text.

The primary focus of Coh-Metrix was originally on discourse cohesion and coherence, but eventually took on a large landscape of language and discourse measures. *Cohesion* is defined as characteristics of the explicit text that play some role in helping the reader mentally connect ideas in the text. *Coherence* is defined as a cognitive representation that reflects the interaction between linguistic/discourse characteristics and world knowledge. Coh-Metrix was designed to move beyond standard readability formulas, such as Flesch-Kincaid Grade Level (Klare, 1974-1975), that rely exclusively on word length and sentence length to define text difficulty. Some CohMetrix measures refer to characteristics of individual words, but the majority scale texts on deeper levels that analyze syntax, referential and semantic cohesion, dimensions of the situation model, and rhetorical composition.

As an illustration of Coh-Metrix, consider the following passage on the drug *pseudoephedrine*. This popular drug helps those who periodically get nasal congestions from allergies, colds, or flu's. However, it once was abused by drug addicts so the drug market had to seek substitutes. The public has wanted to know about the uses, warnings, and side effects of this class of drugs and the associated substitutes. The passage below is on the container of one of the commercial drugs with pseudoephedrine.

> Temporarily relieves nasal congestion due to the common cold, hay fever or other upper respiratory allergies, and nasal congestion associated with sinusitis. Temporarily relieves sinus congestion and pressure. Do not use if you are now taking a prescription monoamine oxidase inhibitor (MAOI) (certain drugs for depression, psychiatric, or emotional conditions, or Parkinson's disease), or for 2 weeks after stopping the MAOI drug.

When we submitted the passage to Coh-Metrix and recorded the scores on a number of metrics, we discovered the passage would challenge most of the public. The Flesch-Kincaid readability score showed a 12-grade reading level, so members of the public who did not manage to graduate from high school would allegedly have trouble reading the container about the use and warnings of pseudoephedrine. The word frequency scores were low for a number of words (monoamine, oxidase, inhibitor), which would create comprehension problems. The conceptual cohesion between successive sentences in the text was intermediate for referential cohesion and low for situation model coherence, whereas there was a high density of logical operators (*and, or, if, then, not*). This profile of measures supports the conclusion that readers will need to have high reasoning and verbal skills to comprehend the passage, so would not effectively reach major segments of the public.

This Coh-Metrix example illustrates the importance of integrating corpus linguistics and computational linguistics into research projects that investigate the cognitive foundations of discourse processing. An understanding of cognition requires a systematic analysis of discourse in the real world and the computational foundations for computing meaning. Researchers also are integrating neuroscience into their research projects, which is an exciting trend that is beyond the scope of this researcher to address (Schmalhofer & Perfetti, 2007). This is yet another sign that discourse cognition has fulfilled its time-honored explicit mission to be interdisciplinary. Nearly every major research project in contemporary discourse cognition involves a serious interdisciplinary collaboration with corpus linguistics, computational linguistics, and/or neuroscience.

The Challenges of Comprehending Technical Texts

Citizens in most cultures are expected to keep up on advances in science, technology, and other technical topics that are relevant to their jobs and daily lives. However, both children and adults have trouble comprehending technical expository text at deep levels even though they are skilled readers. Deep comprehension of technical text is a difficult challenge because the reader has minimal knowledge of the technical terms, key conceptualizations, mental models, and other forms of background knowledge. Even those with relevant background knowledge and general reading skills can struggle.

The challenges of comprehending technical text have been documented in studies that use tests that assess understanding of the situation model. In one study on Newtonian physics (VanLehn et al., 2007), college students were assigned to one of three conditions: (1) work on physics problems with an intelligent computer tutor (called AutoTutor, which will be discussed later), (2) read a textbook on the same content for the same amount of time as the AutoTutor condition, versus (3) read nothing. Before and after training, there was a pretest and a post-test with multiple choice questions that tested their understanding of deep physics knowledge. There were substantial learning gains from AutoTutor, which was positive news for the developers of that tutor. A more surprising outcome, however, was that the posttest scores showed no learning gains from reading the textbook. The scores in the textbook condition did not differ from reading nothing at all. A similar finding was found on the topic of computer literacy and critical thinking on the scientific method (Graesser, Jeon, & Dufty, 2008).

There are many reasons why deep knowledge does not routinely emerge from reading textbooks. The reason cannot be attributed to the use of poor textbooks because the textbooks in all of the above studies were very popular (the physics text was on the 8th edition) or written by accomplished scientists. The reason is more likely to be attributed to readers' taking insufficient time and effort to construct deep and accurate situation models. However, that reason raises another question: Why didn't the readers take the time and effort to accurately comprehend the material? Two reasons are most likely. First, the readers are not able to calibrate their comprehension with sufficient accuracy. Second, they do not use reading strategies that facilitate deep comprehension.

The notion that comprehension calibration in adults is faulty has been confirmed in dozens of studies in discourse cognition (Maki, 1998). Comprehension calibration is measured by computing a correlation between (1) the readers' ratings of how well they comprehend a text and (2) scores on an objective test of text comprehension. The correlation is modest (r = .27), indicating that many readers do not have an accurate sense of when they are comprehending versus not comprehending the material. Discourse psychologists have documented many other findings that support the general conclusion that adult readers routinely settle for a shallow standard of comprehension (Sanford & Sturt, 2002). Readers often miss contradictions in text, unless the premises are prominently juxtaposed in the discourse. They get fascinated with seductive details and interesting tidbits (e.g., *Isaac Newton's mother wanted him to be a farmer*) that distract them from the deep situation model (e.g., *force equal mass times acceleration*). Readers miss faulty presuppositions that obviously clash with prior knowledge. For example, consider the exchange in a classroom below.

Teacher: How many animals of each kind did Moses put on the Ark?

Class (children, in unison): Two.

Teacher: Ah hah. It was Noah, not Moses, who had animals on the Ark.

This *Moses illusion* illustrates that comprehenders may deeply monitor information in the discourse focus, but usually gloss over the presupposed information (Van Oostendorp & Kok, 1990). They adopt the assumption that the speaker/writer expresses only true presuppositions, so this information is not deeply scrutinized. It is frequently the comprehenders with high background knowledge who admit that they are not understanding. This is because it takes knowledge to know what one does not know. Researchers who investigate metacognition (i.e., what people believe about their thinking processes) have documented many of these paradoxes and counterintuitive results (Hacker, Dunlosky, & Graesser, 2009).

The comprehension strategies of readers are usually not pitched at deep standards of comprehension when the text is technical. Therefore, they need to learn better strategies. One of the exciting trends in discourse cognition has been to develop learning environments and interventions to help students learn better comprehension strategies (McNamara, 2007). There are strategies designed to improve: (a) the comprehension of sentences and local text excerpts, (b) the bridging and connecting of text constituents, (c) the grounding of the text to personal experiences and everyday activities, (d) mastery of the rhetorical structure and genre of text, (e) social interaction with experts, tutors, and peers, and (f) processes of question asking, question answering, reflection, and summarization. One strategy, for example, is to encourage students to generate *self-explanations* of the text while reading. This is because self-explanations are known to improve comprehension (Chi, de Leeuw, Chiu, & LaVancher, 1994; McNamara, 2004). That is, as the reader comprehends the text, sentence by sentence, the individual constructs explanations that explain why events occur, why the author expresses something, how ideas are connected, and how the material is related to the reader's prior knowledge and personal experiences. Comprehension strategy training has traditionally occurred in the classroom or through human tutoring, but recently has shifted to computer environments.

Electronic Texts, Discourse, and Multimedia

It seems that we are doing more and more reading on digital displays, whether they are computers, cell phones, PDAs, laptop computers, e-book readers, or video monitors. What we are reading also appears to be more varied than the genres mentioned earlier: emails, instant messages, multimedia displays, and surfing the World Wide Web. The popularity of such gadgets certainly speaks to our desire to keep in contact with other people in addition to completing tasks more efficiently.

Technology is indeed changing the way that we read. More of what we read is presented on a digital display. There are obvious advantages to having discourse presented on a digital device. One advantage is that through e-book and other portable devices, we can have a wealth of discourse available to us 24/7. For example, the *Kindle* system sold by Amazon.com uses WiFi technology to download thousands of books. Another advantage is that the technology can aid the reader in particular levels of comprehension. We can change the font size of text or look up a word's meaning by clicking on a dictionary tab. These would correspond to levels 1 (surface) and 2 (textbase), respectively. There is no reason to stop at these levels. In the near future, there could be short interchanges between the e-book and the reader to assess whether the reader has formed an adequate situation model.

It is too early to tell whether all of our paper-based books will be replaced by e-books. We do know that the cost of books versus eBooks is quickly changing so it is impossible to say which is most economical. Aside from the costs, there is something familiar and soothing about holding a book, and flipping through its pages, especially on a cool rainy day near the fireplace. Some people, especially those with low vision, have difficulty reading digital displays. If the contrast between the letters and background is not just right, then they have trouble focusing and their eyes get tired. Fortunately, vision specialists and researchers in psychology are teaming with private industry to help solve problems arising from reading text presented on digital displays. One solution is called *visual syntactic text formatting*. Instead of presenting text in regular paragraph form, the text is presented this way:

You are reading

this sentence

as formatted

by visual-syntactic text

formatting.

Despite its odd appearance, there is growing evidence that this type of formatting reduces eye strain and increases comprehension (Walker, Schloss, Fletcher, Vogel, Walker, 2005). It is believed to increase comprehension by helping the reader syntactically parse the sentence (level 1) so the reader can spend mental resources on higher levels.

There are other technological means for helping us understand the written word. One is multimedia, where information is presented via different presentation modes (verbal, pictorial) and sensory modes (visual, auditory, and perhaps touch) that use different delivery systems (video, text, simulations, pictures). Information presented through multimedia increases memory and comprehension, provided that it does not produce attention overload. One reason why pictures (photographs, videos) facilitate memory is by increasing the number of ways that the information can be retrieved at a later time. For example, a student who is shown a diagram of a hydrogen atom would show better memory than a student who reads a verbal description of it. The student seeing the diagram would encode both its image and a verbal representation of it ("the nucleus is in the center" "electrons circle the nucleus" etc), whereas the student who only reads the description may not produce a mental image. Therefore, the student seeing the diagram would have a greater chance at retrieving one of the representations at a later time (verbal description or an image).

In addition to memory, multimedia can help readers achieve a deep understanding of the material. Pictures, graphical displays, animation, and videos all have the potential to help readers achieve a deep, complete and accurate situation model of the accompanying text. Imagine reading a text on the inner workings of a computer: the CPU, XOR gates, bits, inputs and outputs, and so on. Readers might have trouble comprehending the complex temporal, causal, spatial, logical, and semantic relations which exists in such a complex domain. However, many of these relations can be conveyed quite easily in a nifty animation, graph, or simulation in which the user can control some parameter (e.g., a logic gate, the size of a memory) and observe the consequences.

Mayer (2005) has extensively studied how multimedia affects comprehension, memory, and learning. He has generated sets of principles that specify how multimedia can increase learning based on how information is selected, organized and integrated. Although multimedia can help comprehension by providing a semantically rich integrated experience, multimedia displays can also interfere with comprehension. For example, a series of static images that convey important functions might actually be better than an animation because the latter might occur too quickly for a viewer to apprehend all of the information (Ainsworth, 2008). In addition, there are split attention effects. This happens in Powerpoint presentations when a speaker is talking at the same time that a word-rich slide is being shown. The viewer's attention is being split between the speaker and the slide and is overloading working memory, so little information is being comprehended. Similar processes occur in multimedia displays if any one modality is overloaded at critical moments.

We are living in the time of eLearning, when much learning, training, and education is being delivered through electronic means. In the old schools, students and co-workers would meet at a specified place and time to be instructed by a human teacher or boss. In the new schools, we are instructed through a tutorial over the web that we can view at home or on an iPhone while riding on the bus. The advocates of eLearning point to a future when learning is more student-centered, active, engaged, not limited to a particular space and time, multimedia-rich, individualized, contextualized in realistic situations, problem-based (versus instruction-based), social, and opportunistic of portable computing devices. The teacher becomes a guide on the side rather than a sage on a stage. The growing use of electronic documents requires a new set of literacy skills, such as evaluating the source of documents on reliability and relevance (Wiley, Goldman, Graesser, Sanchez, Ash & Hemmerich, in press).

The "net" generation seems to be well-suited for such a future because they have grown up with the World Wide Web, instant messaging, cell phones, and video games (Gee, 2003). Indeed, some researchers in education, psychology, discourse processing, and artificial intelligence are teaming with the gaming industry in an effort to create engaging educational games. They have noticed the vast appeal of Massively Multi-player On-line Role-Playing Games (MMORPGs) where thousands of players interact virtually (being represented by "avatars") and simultaneously over the Web. In 2008, there were 10,000,000 subscribers to *World of Warcraft*, a fantasy-based MMORPG. To the dismay of many parents, the average teenage gamer plays video games for around 20 hours per week, far longer than academic homework. Consider the impact of having a deeply engrossing video game that teaches in-depth biology, chemistry, algebra, or political science concepts (Ritterfeld, Cody, & Vorderer, 2009). Most parents would be thrilled if their son or daughter actually looked forward to their homework.

Animated Conversational Agents

Virtual learning environments (educational games, simulations, multi-user virtual environments) are increasingly using animated pedagogical agents. Pedagogical agents are virtual entities that come in a variety of forms (e.g., people, plants, animals, insects, common objects) that can move (e.g., walk, point, gesture), communicate (speak, listen), show personality (e.g., rude, polite), simulate different cognitive states (e.g., learn, forget) and display emotions (e.g., surprise, happiness, skepticism). Because an "agent" might be defined as any entity with a point of view, developers of learning environments have provided animated agents with a number of pedagogical roles or functions: teacher, tutor, student, expert, friend, on-looker, and so on (Baylor & Kim, 2005). In some learning environments, there might be a single animated agent, while in others there might be 2 or 3 or even more. The animated agents can communicate with the human user by asking and answering questions, but the animated agents can also communicate with each other. For example, a 'student agent' can model a task for the human user by holding a conversation with the 'teacher agent.' As one might expect, animated pedagogical agents can be very engaging to interact with, opening up a whole new world of possibilities. Moreover, discourse psychologists have played a central role in building and testing these conversational agents (Graesser, Jeon, & Dufty, 2008; McNamara, 2007; Millis et al., 2009). One powerful way of advancing our understanding of conversation is to build an agent that can successfully communicate with humans and other agents.

Animated agents in learning environments provide a natural and engaging interface for navigating the user (student) toward acquiring a deep understanding of the topic. As mentioned earlier, most readers do not know whether they truly understand what they are reading. Pedagogical agents can be used to challenge the reader's understanding so that incomplete knowledge and misconceptions can be identified, addressed and fixed. One example of a learning environment that has been successful in this endeavor is AutoTutor (Graesser, Jeon, & Dufty, 2008; Van Lehn et al., 2007). AutoTutor is a computerized intelligent tutor that helps students learn by holding a conversation in natural language. The AutoTutor agent poses problems to the user, assesses the quality and completeness of the answer, and guides the user in forming a complete answer by giving short feedback (positive, negative, neutral), hints, prompt queries to elicit a specific word, and pumps ("what else?"). AutoTutor answers some types of student questions and provides summaries at various points in the dialogue. AutoTutor relies on methods in computational linguistics to "understand" the answers given by the human user. AutoTutor has so far been developed for science and technology topics, such as physics, biology, research ethics, computer literacy, research methods, and critical scientific thinking. AutoTutor has been successful in increasing students' comprehension. Based on several studies, students interacting with AutoTutor score nearly a letter grade above various control conditions on tests of comprehension and learning (Graesser et al., 2008). This is roughly double the effect size of most human tutors that students interact with in middle school through university settings.

The framework behind AutoTutor can suit many different types of goals in various learning environments. One learning environment is an educational game called

Operation ARIES! that is currently being produced. ARIES is an acronym for **A**cquiring **R**esearch and **I**nvestigative and **E**valuative **S**kills (Millis, Cai, Graesser, Halpern, & Wallace, 2009). The learning objective of *Operation ARIES!* is for teaching students how to critically evaluate research found in the media. A central key to the storyline behind the game is that aliens from the Aries constellation are discretely trying to confuse human beings about the scientific method by publishing flawed research. In the first level of the game, the human student needs to complete a science course by reading an online textbook about research. The human is given the opportunity to test out of each chapter by taking a "challenge" multiple-choice test, or to wait until after the chapter is read to be tested. At this point of the game, there are two animated agents: Dr. Quinn, who is the teacher, and Glass Tealman, who is a student peer also taking the course.

After the human player answers a multiple choice question, there is a brief conversation among the three agents: the player, Dr. Quinn, and Glass. These are called *trialogs* and are orchestrated by a modified version of AutoTutor. The purpose of the trialogs is to increase student learning through discourse and to help the student regulate their own learning in a fashion that is aligned with their own level of knowledge. When performance on the test questions indicates that the student has low knowledge of the test topic, the human will watch Dr. Quinn engage Glass in a tutorial dialog. There is evidence that low knowledge students are helped by watching others interact (Craig, Sullins, Witherspoon, & Gholson, 2006). When performance indicates moderate knowledge, Dr. Quinn will engage the human student with a brief tutorial dialogue, very similar to the original AutoTutor. Finally, when the human shows high knowledge, she is asked to teach Glass. That is to say, one way to help good students learn even more is to have them teach other students or agents.

Table 1 shows an example trialog in this last condition in which the human teaches a peer agent. The topic is dependent variables, which are the measured outcomes in an experiment. Glass asks Amanda (a real psychology undergraduate) for help defining dependent variable. The system wants Amanda to say that (1) dependent variables are measured and (2) they depend on the independent variable (the variable manipulated by the experimenter). Amanda indicates that she knows (2) but not (1) in her first two contributions. Glass tries to get her to say "measured" by supplying her with a prompt, but she fails to do so, leading Dr. Quinn to interrupt with the correct answer. After a hint is supplied by Glass, Amanda finally articulates the desired idea, namely that the experimenter measures the dependent variable.

One might argue that the dialog is not perfect, and a little clunky. However, these types of interactions encourage students to perform a number of processes that are known to increase comprehension. These include having the student use the desired vocabulary, generate explanations, discriminate relevant from irrelevant information, and address misconceptions. These processes target the surface, textbase and situation model levels of discourse. It should also be noted that Glass needs the human player's help, which will hopefully motivate the human to do well. Here the pragmatic level becomes relevant. That is, the human should be explicit in her contribution since she would assume Glass does not know the information that she does, and hence cannot assume a common ground. Indeed, it is an interesting question of whether humans will assume similar pragmatic stances with animated agents as they do with fellow humans.

A skeptic may object that the conversational agents will ultimately be a disaster when they do not completely understand the human. We suspect this may be true when there are high expectations on precision and common ground. However, this is not the case when the content involves verbal elaboration rather than mathematical operations and when the student has little-to-moderate knowledge about the topic. The agent also has the option of hedging on how well it understands the student (e.g., "I can understand much of what you say but please understand that I am not perfect and cannot be a mindreader"). The fundamental test of the pedagogical value of these conversational agents does not lie in their ability to comprehend perfectly, but rather in their comparisons to humans and in their ability to facilitate learning.

Closing Comments

This chapter has reviewed recent research on discourse processing that takes a cognitive slant. Sufficient progress has been made in the fields of discourse processes, cognitive science, corpus linguistics, and computational linguistics to build detailed models of how discourse is comprehended and produced at multiple levels: the surface code, the textbase, the situation model, genre and rhetorical structure, and pragmatic communication. These models are formulated and tested by collecting empirical data with rigorous scientific methodologies. Many of these levels are sufficiently well specified to automate them on computers. The computational models have evolved to the point of building useful computer technologies, such as essay graders, text analyzers that go well beyond readability formulae, automated conversational tutors, and multiparty interactive games. Researchers in discourse cognition have continued their interdisciplinary tradition of embracing the insights and methodologies of other fields, most recently in corpus linguistics, computational linguistics, education, multimedia, conversational agents, and neuroscience. Given the dramatic changes to the field that have occurred during the last decade, it is difficult to forecast what breakthroughs will emerge during the next 10 years. However, the interdisciplinary stance of those in discourse cognition gives us every reason to remain optimistic.

Further Reading

The following is a list of books that the interested reader may learn more about the topics touched upon in this chapter.

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

This work provides a nice overview of contemporary work on discourse processes, including literary, educational, spoken, written, and computer-mediated discourse. The book also presents different theories, perspectives and methodologies used by a set of interdisciplinary researchers.

Schmalhofer, P. & Perfetti, C. (2007)(Eds.). *Higher level language processes in the brain: Inferences and comprehension processes*. Mahwah, NJ: Erlbaum.

The book elaborates on the different levels of comprehension that begin this chapter, in particular the textbase and the situation model. It also emphasizes recent research on the underlying neural representations that support them.

McNamara, D.S. (2007)(Ed.). *Reading strategies*. Mahwah, NJ: Erlbaum.

In order to gain deep understanding of tough material, readers use several different types of reading strategies, such as self-explanation, reviewing, and asking questions. A reader who would like to know more about current theory, interventions and technologies regarding reading strategies will find this book useful.

Spector, J.M., Merrill, M.D., van Merriënboer, J.J.G., & Driscoll, M.P. (2008)(Eds.), *Handbook of research on educational communications and technology*. London: Taylor & Francis.

The handbook examines emerging information and communication technologies (ICT) that are becoming more relevant as we progress through the digital age. It contains sections on historical foundations, instructional and learning strategies, tools and technologies, models of learning, design and development, and methodological issues.

Ritterfeld, U., Cody, M. & Vorderer, P. (2009)(Eds.), *Serious games: Mechanisms and effects*. Mahwah, NJ: Routledge, Taylor and Francis.

Serious games are interactive computer games that are designed to teach skills and content. This book explores psychological and social mechanisms related to serious games, summarizes research on their effectiveness, and discusses their promises and limitations in educational settings.

Bionotes

Dr. Arthur Graesser is a full professor in the Department of Psychology, an adjunct professor in Computer Science, and co-director of the Institute for Intelligent Systems at the University of Memphis. He is the editor of *Journal of Educational Psychology* and has published over 400 articles in journals, books, and conference proceedings, written two books and edited nine books (one being the *Handbook of Discourse Processes*). He has designed, developed, and tested cutting-edge software in learning, language, and discourse technologies, including AutoTutor, Coh-Metrix, HURA

Advisor, SEEK Web Tutor, MetaTutor, Operation ARIES, Question Understanding Aid (QUAID), QUEST, and Point&Query.

Dr. Keith Millis is a full professor in the Psychology Department at Northern Illinois University. He completed a post-doc at Carnegie Mellon University after receiving his Ph.D. in 1989. He has authored several articles and book chapters on discourse comprehension and inference generation. He is the project director of Operation ARIES!, a serious game designed to teach aspects of scientific inquiry.

Author Notes

Acknowledgements go to the funders of our research on reading comprehension and discourse processing, including the National Science Foundation (ITR 0325428, ITR 0325428, REESE 0633918), the Institute for Education Sciences (IES R3056020018-02, R305H050169, R305B070349, R305A080589), and the DoD Multidisciplinary University Research Initiative (MURI) administered by ONR under grant N00014-00-1-0600. Requests for reprints should be sent to Art Graesser, Institute for Intelligent Systems, 365 Innovation Drive, University of Memphis, Memphis, TN 38152.

Table 1. Example trialog. Glass = animated student, Dr. Quinn = teacher, Amanda = human student.

GLASS: I think I'm confusing independent and dependent variables. Amanda, what is a dependent variable again?

AMANDA: A dependent variable is the variable that is affected by the independent variable.

GLASS: Thinking about dependent variable variables make me think about cause and effect relationships for some reason.

AMANDA: the manipulation of the independent variable is the cause of the dependent variable.

GLASS: I know the dependent variable isn't the variable that's manipulated. So, that means it must be the variable that is what?

AMANDA: controlled.

DR. QUINN: measured

GLASS: Dependent variables have something to do with the data.

AMANDA: they are what the experimenter measures the independent variable against.

DR. QUINN: Brilliant!

GLASS: I got it. A dependent variable is the outcome that gets measured. These measurements depend on the groups that make up the independent variable.

References

Ainsworth, S. (2008). How do animations influence learning? In D.H. Robinson and G. Schraw (Eds.), *Recent innovations on educational technologies that facilitate student learning*. Charlotte, NC: Information Age Publishing.

Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education, 15*, 95-115.

Biber, D. (1988). *Variation across speech and writing*. Cambridge, England: Cambridge University Press.

Chi, M.T.H., de Leeuw, N., Chiu, M. & LaVancher, C. (1994) Eliciting self-explanation improves understanding. *Cognitive Science, 18*, 439-477.

Craig, S. D., Sullins, J., Witherspoon, A. & Gholson, B. (2006). Deep-level reasoning questions effect: The role of dialog and deep-level reasoning questions during vicarious learning. *Cognition and Instruction, 24,* 565-591.

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

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

Graesser, A. C., Cai, Z., Louwerse, M., & Daniel, F. (2006). Question Understanding Aid (QUAID): A web facility that helps survey methodologists improve the comprehensibility of questions. *Public Opinion Quarterly*, *70*, 3–22.

Graesser, A. C., Jeon, M., & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes*, *45*, 298–322.

Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory and Cognition*, *33*, 1235–1247.

Graesser, A.C., McNamara, D.S., Louwerse, M.M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers, 36*, 193-202.

Graesser, A.C., Millis, K.K., Zwaan, R.A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*, 163-189.

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-95.

Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General*, *114,* 357-374.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.) (in press). *Handbook of metacognition in education*. Mahwah, NJ: Erlbaum/Taylor & Francis.

Jurafsky, D., & Martin, J.H. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Boston: Allyn & Bacon.

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

Klare, G.R. (1974-1975). Assessing readability. *Reading Research Quarterly, 10,* 62-102.

Landauer, T.K., Laham, D., & Foltz, P.W. (2003). Automatic scoring and automation of essays with the Intelligent Essay Assessor. In M.D. Shermis and J. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspective* (pp. 87-112). Mahwah, NJ: Erlbaum.

Lea, R.B., Rapp, D.N., Elfenbein, A. Mitchel, A., & Romine, R.S. (2008). Sweet silent thought: Alliteration and resonance in poetry comprehension. *Psychological Science, 19*, 709-715.

Maki, R.H. (1998). Test predictions over text material. In D.J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.). *Metacognition in educational theory and practice* (pp. 117-144), Mahwah, NJ: Erlbaum.

Mayer, R.E. (2005). *Multimedia Learning*. Cambridge, MA: Cambridge University Press.

McNamara, D.S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 1-30.

McNamara, D.S. (2007)(Ed.). *Reading strategies*. Mahwah, NJ: Erlbaum.

McNamara, D.S., Levinstein, I.B. & Boonthum, C. (2004). iSTART: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, and Computers, 36*, 222-233.

Millis, K., Cai, Z., Graesser, A., Halpern, D. & Wallace, P. (2009). Learning Scientific Inquiry by Asking Questions in an Educational Game. In T. Bastiaens et al. (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009* (pp. 2951-2956). Chesapeake, VA: AACE.

Millis, K.K., King, A., & Kim, J. (2001). Updating situation models from descriptive texts: A test of the situational operator model. *Discourse Processes, 30*, 201-236.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*, 1200-1210.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count*. Austin, TX: LIWC.net (www.liwc.net).

Pickering, M., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue'. *Behavioral and Brain Sciences, 27*, 169-226.

Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Ritterfeld, U., Cody, M. & Vorderer, P. (2009)(Eds.), *Serious games: Mechanisms and effects*. Mahwah, NJ: Routledge, Taylor and Francis.

Sanford, A. J. & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences, 6*, 382-386.

Schmalhofer, P. & Perfetti, C. (2007)(Eds.). *Higher level language processes in the brain: Inferences and comprehension processes*. Mahwah, NJ: Erlbaum.

Van Dijk, T.A. (2008). *Discourse and context: A socio-cognitive approach*. New York: Cambridge University Press.

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

VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A., & Rose, C.P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31*, 3-62.

Van Oostendorp, H., & Kok, I. (1990). Failing to notice errors in sentences. *Language and Cognitive Processes, 5*, 105-113.

Walker, S., Schloss, P., Fletcher, C.R, Vogel, C.A., & Walker, R.C. (2005). Visual-Syntactic Text Formatting: A new method to enhance online reading. *Reading Online, 8(6)*. Posted May 2005. International Reading Association, Inc. ISSN 1096-1232.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in Internet science inquiry tasks. *American Educational Research Journal, 46,* 1060-1106.

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