CHAPTER 7

Discourse Analysis

Written Text

Susan R. Goldman
Jennifer Wiley

Written text can be approached from a variety of disciplinary perspectives and purposes. In the context of this volume, we approach discourse analysis of written text by asking why written texts are of concern to literacy researchers, why they need to be analyzed, and how the texts and their analyses inform theoretical and empirical research in literacy. Written texts are of concern to literacy researchers because the ability to read and understand them is definitional to literacy, at least in Western culture. But written text is far from monolithic. There are any number of written text genres, differentiated by their purpose or function as well as their structure or form (e.g., narrative, poetic, persuasive, informative). And within a genre texts vary in both their form and their content. A primary goal of the analysis of written text is to describe structure and content. It is important to do so because well-established empirical findings indicate that structure and content affect how readers read, understand, remember, and learn from written texts (Goldman, 1997; Goldman & Rakestraw, 2000; Hiebert; Englert, & Brennan, 1983; Lorch, 1989). The discourse analysis of written text provides a method for systematically describing texts that students read as well as those they write.

To get a sense of what we might want our analysis of written text to capture, read over the following example texts 1 and 2. They each deal with the topic of river ecosystems and were written to introduce young adolescents (11- to 13-year-olds) to the ideas of interdependence and environmental pollution. As you read them, think about how you might describe their characteristics, their similarities, and their differences. (Do this now before reading on in the chapter.)

TEXT 1: THE TUPELO RIVER MYSTERY

In the spring of 1999, a new nuclear power plant in the town of Bregville went into operation. Local citizens and environmentalists were concerned because the power plant was located along a branch of the Tupelo River. They knew that a river is a very fragile ecosystem. The word ecosystem connects the idea of evo, which means a habitat or environment, with that of system, which means a set of relationships. Environmentalists were worried that the nuclear power plant would upset the balance of relationships among everything in the Tupelo River ecosystem.

There are many parts of the Tupelo River's ecosystem. Some are organisms such as fish, insects, animals, and plants. They live in and around the river. River systems also contain nonliving things such as water, the rocks on the bottom, and the mud on the banks of the river. The most important concept about any ecosystem, including the Tupelo River's, is that everything depends on everything else. A change in one part of the ecosystem could introduce changes in all parts of the ecosystem. There was concern that the operation of the nuclear power plant would introduce changes.

TEXT 2: ECOSYSTEMS

Any group of living and nonliving things interacting with each other can be considered as an ecosystem. Within each ecosystem, there are habitats that may vary in size. A habitat is the place where a population lives. A population is a group of living organisms of the same kind living in the same place at the same time. All of the populations interact and form a community. The community of living things interacts with the nonliving world around it to form the ecosystem. The living members of a river's system are the plants and animals that live around and within the river—and they are all connected to each other in what is called a food web.

Anything that is added to the water of our aquatic ecosystems that is not a normal part of the systems, and that should not be there, is a type of water pollution. There are many sources of water pollution. Some types of pollution can be traced directly to a particular spot, such as a factory or industrial plant. These sources of water pollution are easier to control because the actual point where the pollution is being added to the water can be identified.

Here are some of the things you might have come up with to describe the characteristics and the differences.

1. Text 1 is a story about a town; text 2 is more straightforward "content."
2. They both define ecosystem but they do it differently.
3. They are both about the same length and have the same number of paragraphs.
4. Text 2 talks about water pollution in general; text 1 talks about the impact of the nuclear power plant but doesn't say it is a potential source of pollution.
5. Text 2 talks about the food web, but text 1 just says that things in an ecosystem depend on each other.
We suspect that as you engaged in the analysis and comparison process, you were unsure of precisely how to describe what you were noticing. Discourse analysis provides a means to engage more systematically in the descriptive analysis and comparison of written texts. The particular perspective we take in this chapter assumes that discourse analysis is informed by knowledge of the content domains that the text is about. In other words, the method is not "content free."

WHAT IS THE METHOD?

Discourse analysis of written text is a method for describing the ideas and the relations among the ideas that are present in a text. The method draws on work in a variety of disciplines, including rhetoric, text linguistics, and psychology. To see these disciplines as interrelated and analytical, the structure and content of the text encodes ideas and the relations among the ideas. In describing these relations, it is important to initially define the genre to which the text belongs because structures differ across genre. For example, narrative stories differ from persuasive essays; news articles have a different form than editorials; and fiction texts have different structures than nonfiction. Differences in structure imply different relations among the ideas in the texts, especially at the global level. With respect to the differences between the example texts 1 and 2, the first difference we listed reflects just such genre differences.

The global level of relations is often referred to as the rhetorical structure of the text. Different rhetorical structures are appropriate for different genres. For example, the rhetorical structure of many stories involves the occurrence of a problem that protagonists attempt to resolve and, in "happy ending" stories, are successful at resolving. The story typically consists of a series of episodes that are causally related. A common form of episode relations occurs when there are a number of preconditions that must be met in order to resolve the overall problem. In other narrative rhetorical structures, the interepisode relations may be temporal, with one episode succeeding another but with goals that are not particularly related. (For further information about narrative structure, refer to Bamberg, 1997; Bloom, 2003; Mandler & Johnson, 1977; McCabe, 1997; Stein & Albro, 1997; Stein & Glenn, 1979; Trabasso & van den Broek, 1985.)

The global structure of nonfiction texts contrasts with those of fictional texts, and there is less agreement about the rhetorical structures that apply. Meyer (1985) proposed a set of five "top-level" rhetorical structures in an attempt to systematize the structure of the major genre of expository texts: collection or list, description, causal, comparative, and problem/solution (see also Weaver & Kintsch, 1991). These rhetorical structures may be signaled by particular words or phrases but do not have to be. Example text 2 is a description. Example text 1 is a bit more complicated, but we think it may be best described as a nonfiction narrative perhaps similar to a feature article in a news magazine or newspaper.

The analysis of written text is also concerned with understanding the local relations among the ideas conveyed in a text (i.e., relations among information in sentences occurring relatively close together in the text). It is precisely the relations among ideas that define the coherence of the text and make it more than the sum of its parts. Indeed, Sanders, Spoorien, and Noordman (1992) defined coherence relations as the "aspect of meaning of two or more discourse segments that cannot be described in terms of the meaning of the segments in isolation. In other words, it is because of this coherence relation that the meaning of two discourse segments is more than the sum of its parts" (p. 2). There are a number of ways coherence relations are established. Halliday and Hasan (1976) identified four primary logical connective devices: temporal, spatial, informational, and causal; as in antecedent and consequent events; and adversative, as in the juxtaposition of contradictory information.

A number of discourse markers (also called linguistic cues) are relevant to understanding local relations among ideas. A specific class of discourse markers, called connectives, express, signal, or cue the underlying conceptual coherence relations. Examples of connectives are because, furthermore, and however. Other kinds of signaling devices (Lorch, 1989) signal the relationships of sentences to paragraphs and of paragraphs to one another and to the overall theme of the text.

Lorch (1989) distinguished signaling devices from those aspects of a text that communicate the semantic content: Signals emphasize particular aspects of content or structure, but they do not add content (Lorch, 1989). Signaling devices help readers pick out what is selectively arranged to and how to differentiate the importance of different information in the text (Goldman & Durán, 1988; Goldman & Saul, 1990; Guthrie, 1988; Lorch & Chen, 1986; Lorch & Lorch, 1996). Lorch identified a variety of signaling devices used in expository prose, including titles, headings, and subheadings; repetition of content to emphasize, preview, or summarize; function indicators (pointer words such as thus, pointer phrases such as in summary, summarizing what has been said); relevance indicators (let me stress that); enumeration devices; and typographical cues (underlining, boldface, and spatial layout such as indenting and centering).

A large body of empirical research indicates that some genres and content structures are more difficult for readers to understand than others. For example, knowledge of narrative structure generally appears earlier than knowledge of expository structure (Crowhurst, 1990; Englehard, Gordon,
& Gabrielson, 1992; Grabe, 2002; Langer, 1986; Scott & Windsor, 2000; Tolchinsky, Johansson, & Zamora, 2002). Of course, this may be a result of typical instructional practices of concentrating on narrative text in primary grades and not introducing expository forms until later (Duke, 2000; Pappas, 1993). When readers have knowledge of informational content structures and use that information to guide their processing of text, their understanding and memory are better than in the absence of such knowledge (Carrell, 1992; Engler & Hiebert, 1984; Garner et al., 1986; McGee, 1982; Meyer, Brandt, & Bluth, 1980; Taylor & Beach, 1984; Taylor & Samuels, 1983). Although literacy researchers have always been aware of genre, content, and structure differences among texts, historically there was far less recognition of the role that such differences might play in outcomes for readers and the implications for literacy education. This is no longer the case. Thus, it is critical for literacy educators to describe and understand the content and structure of written texts, the demands they place on readers, and the relationship between written texts students read and those they write.

THE ANALYSIS OF WRITTEN TEXT

The goal of analyzing written text is to arrive at systematic descriptions that provide a basis for comparing written texts with one another. But what kind of description? There are a number of issues to consider in answering this question, the most important of which is what questions the researcher wishes to address with the analysis. If, for example, interest is in how much of the text a student remembers exactly as it was presented, counting the number of words that can be reproduced in correct order and position might be all that is necessary. Of course, even word counting can become more complex (e.g., if location in the text—beginning, middle, end—was important). Indeed, early work on memory for written text used the word-counting method (Clark, 1940, cited in Carrell, 1972; Gomulicki, 1956; King, 1960, 1961, cited in Carrell, 1972). Word counting might also be useful if interest is in a simple measure of fluency of producing written text.

A fairly common reason for wanting a systematic description of written text is to determine the reading difficulty of a passage, referred to as passage readability. Traditional readability formulas such as the Flesch Reading Ease index make extensive use of word complexity (assessed via number of syllables), number of sentences, and number of sentences per 100 words (see Klaire, 1974-1975, for review). However, readability formulas do not correspond well with how easy the text makes it to understand the concepts and ideas. One reason for this is that making explicit the local relation between ideas sometimes results in longer sentences. Consider the contrast between examples 1 and 1a:

1. Torrential rains fell on Saturday. The road collapsed.
1a. The road collapsed because of the torrential rains that fell on Saturday.

Even though example 1a is a longer sentence, readers find it easier to understand the causal connection than in example 1 (Pearson, 1974-1975).

Traditional readability formulas also fail to consider the familiarity of the story elements in the passage. More recently, the Lexile has been used to assess readability (Stenner & Wright, 2002). The Lexile determines readability using sentence length and word frequency. Word frequency is a rough index of concept familiarity and is being widely used. (For additional information, see www.lexile.com.) In addition to underestimating the impact of readers’ knowledge of the passage topic, readability formulas do not take into account the text as a whole. That is, texts have global structure that transcends the individual words and sentences. It is perfectly obvious, for example, that the meaning of a text is entirely different when we reorder the sentences.

To illustrate the limitations of text descriptions that are based only on readability formulas, we need only refer back to the example texts 1 and 2. A number of the differences we noted would have been masked if all we had done was look at word counts and readability indices. Comparing example texts 1 and 2, the relevant data are, respectively, 190 versus 201 words; 12 versus 11 sentences; 15.8 versus 18.2 words per sentence; 6 versus 5.5 sentences per paragraph; 47.3 versus 58 for the Flesch Reading Ease; 10.5 versus 9.6 for the Flesch-Kincaid grade level; and 970 versus 1030 for the Lexile measure. On the basis of these data the two texts are hard to distinguish. Thus, readability formulas provide only a rough descriptive and comparative index and miss many of the important characteristics of texts.

Typically, it is the ideas in text and their interrelations that are of interest to the literacy researcher. Some ideas can be conveyed in a word or two; others require several sentences and paragraphs. How then do we systematically describe the ideational content and structure of written text?

The most commonly accepted unit of analysis among literacy researchers is some form of the proposition, a construct that appeared in the work of a number of linguists, psychologists, and computer scientists during the 1970s (e.g., Anderson, 1976; Fillmore, 1968; Fredericksen, 1973; Kieras, 1981; Kintsch, 1974; Kintsch & van Dijk, 1977; Meyer, 1975; Schank, 1972; van Dijk, 1972). A widely used formulation, and the one we favor, is that of Kintsch and van Dijk (1978). Specifically, van Dijk and Kintsch (1983) defined the proposition as a theoretical unit that corresponds roughly to the meaning of a clause. A whole text is represented by organizing the set of propositions derived from the clauses and sentences in the text. Different forms of organizing the propositions include hierarchical lists, semantic networks, and procedural networks. The proposition itself consists of the simple concepts or atomic propositions and the propositional scheme...
(or complex proposition) into which the atomic propositions are arranged. The propositional scheme consists of a predicate and one or more arguments. Predicates are main verbs of clauses or connectives between clauses. Arguments have functional roles with respect to the predicate (e.g., agent, patient, and object location) or can be embedded propositional schemes. A proposition refers to a state, an event, or an action. The psychological plausibility of clause-level propositional schemes has been demonstrated in a variety of research studies, although we do not review them here (Anderson & Bower, 1973; Gruczer, 1981; Kintsch & Keenan, 1973; Ratcliff & McKoon, 1978).

In Figure 7.1, we present a schematic of the propositional scheme for the first sentence in example text 1: In the spring of 1999, a new nuclear power plant in the town of Bregsville went into operation. To generate the propositional representation shown in Figure 7.1, we analyze its elements and decide whether it is an event or a state. As shown in Figure 7.1, the decision we reach is that it is an EVENT. The decision is based on our understanding that the words went into operation mark a change in state of some entity, in this case a new nuclear power plant. Because it is an inanimate entity, we refer to the plant as an Object. We also know that events occur in time and space, and indeed the first sentence provides us with this information. Time and place are indicators of the circumstances of the event. The structure of the propositional scheme conveys these relations. But the schematic in Figure 7.1 does not accurately reflect the atomic level of analysis. The atomic level, resulting from further analysis of the phrases in the schematic, is shown in the lower portion of Figure 7.1. The atomic level can be considered the basic, core level of meaning, and although it may be useful from linguistic or philosophical perspectives to capture this level, it is often too detailed for purposes of looking at understanding and learning. For our purposes, we typically represent propositions at the more molar level shown in Figure 7.1. Figure 7.2 shows the same sentence represented as a semantic network. We have left the links unlabeled, but they could be labeled with object, time, and location.

Although it is useful for didactic purposes to work out several sentences at the level of detail shown in Figure 7.1, it is often impractical to draw such structures for every sentence in a written text. The typical way in which we create the propositions of a text is in a list format as shown here.

i. Went into operation [OBJECT: new nuclear power plant, TIME: ii, PLACE: iii]
   ii. In the spring of 1999
   iii. In the town of Bregsville

We refer to the proposition labeled i as the predicate proposition; ii and iii specify the time and place of the main action. We use indentation to show that propositions ii and iii are subordinate to the predicate proposition. Depending on the discourse analyst, there might be additional propositions reflecting the adjectival modifiers of plant—for example, is TYPEOF plant: nuclear power; is MOD plant: new. These would also be indented to show that they are subordinate to the main action. The level of detail needed in the proposition specification depends on the questions the researcher is addressing. The next sentence in the paragraph would be similarly represented, as shown in Figure 7.3. In sentence 2 there are two clauses that are causally related, shown in the figure as a BECAUSE node (shaded and rectangular). The argument power plant is common to sentences 1 and 2; this overlap is shown in Figure 7.3 by having the node, shaded for emphasis, linked to the predicate of each sentence. Similarly, sentences 2 and 3 share an argument (local citizens and environmentalists). Understanding
In spring of 1999

In town of Brgsville

went into operation

new nuclear power plant

Identification of the clauses in the first two sentences:
1. In the spring of 1999, a new nuclear power plant in the town of Brgsville went into operation.
2.1 Local citizens and environmentalists were concerned
2.2 because
2.3 the power plant was located along a branch of the Tupelo River.
3.1 They knew 3.2
3.2 that a river is a very fragile ecosystem.

Shorthand of the Propositional Schemata in the first three sentences:
1.1 went into operation (OBJ: new nuclear power plant, TIME: 1.2, PLACE: 1.3)
1.2 in the spring of 1999
1.3 in the town of Brgsville
2.1 were concerned (AGENTS: local citizens and environmentalists)
2.2 because (2.1, 2.3)
2.3 was located (OBJ: power plant, LOC: 2.4)
2.4 LOC: 2.3, along a branch of the Tupelo River
3.1 They knew 3.2
3.2 is (OBJ: river, OBJ: ecosystem)
3.3 very fragile (ecosystem)

If we carried out the process demonstrated in Figures 7.2 and 7.3 for the remainder of the sentences in the paragraph and passage, the result would be an organized network or list of propositions connected on the basis of the meaning relations among the ideas. Organizing the individual clauses into a network or list depends on the semantic relationships among the ideas, often producing a hierarchically organized structure of super- and subordinated ideas. The title and first several sentences usually establish concepts that are subsequently repeated throughout the passage as more information about them is provided. Subsequent incoming information “attaches” to these concepts, creating the subordinate or supporting relation, as illustrated in Figures 7.1, 7.2, and 7.3. Sentences that have many subsequent sentences connected to them take on more superordinate, thematic status in the passage (e.g., Goldman et al., 1996; Kintsch & van Dijk, 1978; Meyer, 1975).

Texts in which the individual clauses are explicitly connected have high cohesiveness (Halliday & Hasan, 1976). Example 1a, previously discussed, is a highly cohesive sentence. Coherence is reflected in the connectedness among propositions in readers’ representations. In example 1, readers might infer the causal relation that is explicit in example 1a, in which case the representation would be similarly coherent for examples 1 and 1a. Similarly, in Figure 7.3, we noted that the reader needs to make an anaphoric refer-
ence to resolve the referent for the pronoun they in sentence 3. If readers do not make that connection, they might make one between Tupelo River in sentence 2 and river in sentence 3, an inference based on the specific river being an instance or example of the general class of rivers. If readers failed to make either the anaphoric or the general-to-specific inference, the two sentences would be unconnected and a gap in local coherence would exist. Lack of coherence in readers' representations of text make it difficult to understand the meaning of the text as a whole; information remains in bits and pieces, so to speak.

To foster the development of coherent representations, texts often supply cues (cohesive devices). This should make it easier to establish coherence but only for those readers who understand how to interpret the cues (Goldman & rakestraw, 2000). Texts that have few cohesive devices require that readers work to fill in the gaps among ideas. Doing so frequently requires that readers have requisite prior knowledge of the content domain or knowledge of the genre of the text. If they do not and the text has few cohesive devices, understanding is often not very good. As might be expected, comprehension by readers who know little about the topic of a text is aided by cohesive devices. However, researchers have found that readers who have high content knowledge achieve deeper understanding of a text in which the cohesion is not obvious compared with those in which cohesion is explicit (Kintsch, 1990; Loxterman, Beck, & McKeown, 1994; McNamara, Kintsch, Butler-Songer, & Kintsch, 1996; Voss & Silfies, 1996). Discourse analysis permits researchers to compare texts in terms of propositional structure and the ease with which connections across propositions, hence coherence, can be achieved.

The interaction with prior knowledge notwithstanding, text cohesion is an important predictor of text difficulty that is not taken into account by traditional readability indices even when augmented by word frequency measures as in the Lexile method. In response to the need to develop indices of text cohesion, Graesser, McNamara, Louwerse, and Cai (2004) have developed a computational approach called Coh-Metrix. Coh-Metrix calculates indices of predictors of text difficulty at the surface, propositional, representational, and genre/rhetorical levels (Graesser & McNamara, in press). Indices that have been found to distinguish between high- and low-cohesion texts include coreferential noun overlap, sentence-to-sentence meaning overlap, and causal explicitness (McNamara, Louwerse, McCarthy, & Graesser, 2010). In the same study, McNamara and colleagues (2010) reported that surface code indices by themselves did not reliably discriminate between high- and low-cohesion texts. The research around the development of Coh-Metrix empirically demonstrates that we need to account multiple levels of discourse in determining text difficulty. There is a public version of the Coh-Metrix tool (cohmetrix.memphis.edu, version 2) that we encourage interested readers to explore.

**The Role of Prior Knowledge**

Why should it be the case that readers with high knowledge of the topic achieve deeper understanding when they have to create coherence? Answering this question involves expanding our discussion of text representation to include an interpretive level. Researchers in the early 1990s recognized a distinction between creating a representation of the text itself and integrating the information in the text with prior knowledge, thereby constructing a representation of the situation described by the text. The representation of the text itself, labeled the textbase, is what we have been constructing thus far (see Figure 7.3). The representation of the situation is referred to as a mental model (Johnson-Laird, 1983) or situation model (van Dijk & Kintsch, 1983) and reflects readers' interpretations of the meaning of the text (Perfetti, 1989), constructed by integrating prior knowledge with the ideas in the textbase. Interpretations come about through elaborative, explanatory, and evaluative processes. To conclude that learning has occurred, evidence of a situation model representation is needed (Kintsch et al., 1993; McNamara et al., 1996). Assessment of learning thus involves going beyond reproductive or recognition memory; readers need to demonstrate that they have formed a coherent model of the situation described in the text. This can be demonstrated in a variety of ways, such as by applying the information to a new situation (Wiley & Voss, 1999), verifying inferences that go beyond the information presented in the text (Wiley et al., 2009), providing verbal explanations or drawings that illustrate how something works (e.g., Chi, De Leeuw, Chiu, & Lavancher, 1994), or successfully carrying out a procedure (Mills, Diehl, Birkmire, & Mou, 1995).

**Discourse Analysis of Texts Produced by Readers and Learners**

Written discourse is not only a medium of information presentation. It is also a window into the mental model of the learner. Frequently, readers and learners are asked to produce text, sometimes to demonstrate what they have learned from reading a particular textbook selection, passage, or set of passages. These learner-produced texts can be analyzed and compared with what learners read or were presumed to have read. Other times learners produce written discourse "spontaneously" and in the absence of some just read "stimulus" material. In these situations, the written discourse communicates writers' content knowledge, beliefs, feelings, and command of the language. In the discussion that follows, we focus on learner-produced writing in the context of presented texts and only briefly discuss spontaneous writing.

**Writing in Response to Presented Text**

When learners produce essay responses to open-ended questions, their writing samples can be analyzed with respect to "how close" they are to
the presented text and how much they go beyond what was "in" the text through inferences and elaborations. We can look for evidence of coherence in the mental model and of appropriate integration of new and prior knowledge, both important benchmarks of understanding (Coté & Goldman, 1999; Coté, Goldman, & Saul, 1998). As such, they act as a window into the internal representations that readers have constructed. By comparing texts learners write with those they have read, it is possible to characterize comprehension and understanding. For example, we might want to know the degree to which they have accurately understood the meaning of the text or how they have interpreted the information, as reflected in elaborations and explanations. To facilitate such analyses, we want to conduct discourse analyses of the written texts produced by readers and learners and do so in a form that is compatible with the representation of the material that they read and from which they learn.

Responses to Specific Prompts and Questions

Consider the three student-produced written texts provided in Table 7.1. These are the responses of three seventh-grade students who had read "The Tupelo River Mystery" text in its entirety. They were asked to respond to the specific probe item, "Explain the idea of an ecosystem." The researchers (Goldman et al., 2003; Oney et al., 2003) used this type of probe because they were interested in what students had understood about ecosystems. They were not interested in learners' memory for the story per se, rather than asking for recall of what they remembered from "The Tupelo River Mystery," the researchers gave students a targeted prompt.

To understand the students' writing samples, it is important to know the gist of "The Tupelo River Mystery" beyond the two paragraphs in the example at the beginning of the paper. (Full text is available at lisd.pech. wic.edu/districtriver_7th.doc.) Briefly, scientists decided to monitor the river to see whether any changes in the ecosystem occurred. Over a 10-month period they found that the mayfly population had declined by over 70% and the rainbow darter by 50%, but the Elodea plant population had increased by 20%. This was explained in terms of a food chain in which the mayflies were central, serving as food for the darter and eating the Elodea. The cause of the food chain disruption was a 5-degree increase in the water temperature, resulting in less dissolved oxygen and the negative impact on the mayfly population. The scientists traced the rise in temperature to thermal pollution from the nuclear power plant.

We noticed several things about the seventh graders' responses. (Before reading on, you might want to take the researcher's role and read over the responses, thinking about similarities and differences among them and between each and the passage.) First, the responses differ in length; however, longer is not necessarily better. Student 3 has the shortest response but perhaps the best explanation of an ecosystem. Student 2 likens an ecosystem to the food chain and includes a number of specifics about how the death of one animal or plant will affect what happens to others. Student 2 seems to have accurate details but lacks the more interpretive understanding that student 3 seems to have achieved. Finally, student 1 seems to have bits and pieces of what the text contained, but they are not connected in ways that show the interrelationships. If we were to paraphrase student 1's response, it might go something like the following: "An ecosystem is a cycle in which insects are important. They eat plants and are a food source for certain fish." The more general notion of an ecosystem is missing. This contrasts dramatically with student 3, who uses the specifics in the story to exemplify the general principles of an ecosystem.

If we want our analyses of these data to be capable of revealing the differences among these three kinds of responses, how are we to accomplish this? We need to specify the conceptual elements of an ecosystem as they are communicated in the text and look for correspondences in the students' responses. We start by dividing the responses into clauses, just as we did the text, and ask whether these clauses are meaning preserving of propositions from the passage. We also look at how the relations among the con-
cepts are expressed and whether they are organized in a way that conveys a coherent situation or mental model of an ecosystem. This is the advantage of student 3’s response over those of the other two. Student 3 has inferred a set of core principles about ecosystems based on the information in the text ("An ecosystem is a relational environmental system. In an ecosystem, everything is related to one another. If one thing is removed, other things are effected."). These principles are then exemplified with the specific species and causal chain (water temperature increase → decrease in mayfly → decrease in fish and increase in plant food). Student 3 seems to have used information from throughout the whole text in developing a mental model. In contrast, student 1 appears to be largely focused on the information contained in the two paragraphs in the text about the food chain. Like student 1, student 2 seems to draw on these paragraphs but the response is less tied to the specifics of insects and fish, suggesting more integration with prior knowledge of endangerment and interdependence in an ecosystem. These ideas are similar to the principles student 3 expresses, yet student 2 does not abstract them beyond specific animals. We might conclude that student 2 has a more concrete or specific situation model than student 3.

Specific prompts such as the one illustrated for "The Tupelo River Mystery" and other kinds of open-ended questions are one way to elicit information from learners and are quite similar to the kinds of questions teachers might pose to students on end-of-unit classroom tests for purposes of assessing learning. The recall task is another method that researchers commonly use to elicit text from students.

Writing in Response to Presented Text: Recall

The recall (or retell) task asks students to write down (or say) what they remember from the text they have heard or read. This task is quite commonly used with many text genres. Here we focus on recall of an informational passage on metabolism typical of those used by many researchers. In research with passages of this genre, subjects typically recall about 25 to 40% of the predicate propositions from the textbase. In crediting students with recalling propositions, researchers need to decide whether to adopt stringent or loosen criteria. Using stringent criteria means that for credit to be given the students’ wording must be very close to the passage wording. More commonly, loosen criteria are used so that credit for inclusion of a proposition is based on comparability of meaning across the passage and students’ productions. Comparability of meaning is usually based on the core idea of the predicate proposition.

The "Metabolism" passage shown in Table 7.2 has been used in research with children ranging in age from 8 to 12 years (Côté et al., 1998; Côté & Goldman, 1999). The "Metabolism" passage, along with several others that were used in the same experiment, was written to begin in the style of a feature article in a newsmagazine and to conform to a specific content structure (Côté et al., 1998; Côté & Goldman, 1999).

For example, in the “Metabolism” passage, the first paragraph reported a new scientific invention, which we hoped would provide students with a reason for wanting to learn more about metabolism. From there, the passage structure is typical of the informational text genre. The second paragraph provided definitional information on metabolism; the next four paragraphs provided information on “four factors that affect metabolic rate.” The final paragraph related back to the first, stating how the new invention would measure metabolism. For the most part, each sentence contained one predicate proposition, and in cases where there were two clauses in a sentence, we focused on the predicate that introduced new information. (For example, in sentence 4, Different people have different metabolic rates that indicate how easily they can produce energy, the new information and the main predicate we focused on was the underlined portion because the information in the second clause had already been given in the passage.)

Also in Table 7.2 are two examples that are representative of recall protocols obtained when students read the passage and were subsequently asked to recall “everything you remember.” Several differences between the two recalls and between the recalls and the passage are evident. First, both recalls leave out much of the information in the passage. Student 1 wrote everything in a single paragraph, and student 2 created a list, using up and down arrows, which we assumed meant increase and decrease, respectively. What about content? Both students have changed much of the wording from the exact wording of the passage. This is fairly typical, and researchers often adopt the meaning-preserving criterion: Does the statement in the recall capture the meaning that was in the passage? If so, the student is credited with having remembered the information. By the meaning-preserving criterion, each student would be credited with remembering the four factors. In addition, student 1 included information from the passage to convey the impact the factor has on metabolism. In contrast, student 2 just indicated the relationship. Notice, however, that the way student 1 expressed these relationships is not identical to the propositions in the text. To take just one example, the passage stated:

For example, some foods are hard to digest, such as complex carbohydrates like rice. The body has to work harder to get energy from rice. If a person ate a steady diet of rice, the result would be a higher metabolic rate.

Student 1 wrote:

Rice is hard to digest, so the metabolism would be high for someone on an all-rice diet.
TABLE 7.2. Metabolism Passage and Recalls by Undergraduates

Metabolism:

Customers in many pharmacies may soon be seeing the latest in new devices for the health-conscious. A sports physiologist is developing a metabolism meter, a device that he hopes will measure the human body's ability to produce energy efficiently.

The rate at which the body produces energy is called metabolism. Different people have different metabolic rates that indicate how easily they can produce energy. The same person may have different metabolic rates, depending on the circumstances.

Specifics of animals also have different metabolic rates.

There are several factors that affect metabolic rate. One factor is the type of food a person or animal eats. For example, some foods are hard to digest, such as complex carbohydrates like rice. The body has to work harder to get energy from rice. If a person is on a steady diet of rice, the result would be a higher metabolic rate.

Another factor affecting metabolism is the climate of the environment.

Temperature may cause the metabolism to change. People and animals that live in cold environments need to produce more energy in order to keep warm. Most animals that live in polar regions have high metabolisms. If people move from a warm to a cold climate, their metabolic rates will increase.

Metabolic rate also differs depending on activity level. Changing the level of activity may cause the body to change its metabolism because different activities require different amounts of energy. For example, basketball players use more energy than golfers so their metabolic rates are generally higher.

To some degree, metabolic rate is influenced by genetic inheritance. Children of parents who have high metabolic rates tend to have high metabolic rates also. This is because the body chemistry of the children is a combination of the body chemistry of the parents.

Metabolism is regulated by hormones produced by the thyroid gland, a tiny gland located at the base of the neck. These hormones regulate the behavior of all the cells in the body so that enough energy is produced. The metabolism will work by maintaining hormone levels in the blood.

Student 1

A sports physiologist is trying to measure the metabolic rate by creating the metabolon. Metabolism is affected by many things. It depends on what someone eats. Rice is hard to digest, so the metabolism would be high for someone on an all-rice diet. Climate accounts for metabolism. People in colder climates have higher metabolic rates than people in warm climates. Genetics is responsible for metabolism too. If a child's parents both have high metabolic rates, the child will have a high metabolic rate too. Activity changes the metabolism. A golfer will have a lower metabolic rate than a basketball player because golf does not require as much energy as basketball.

Student 2

Metabolism: Deals with ability to body to produce energy

Influenced by 4 things:
1. Diet: harder food to digest = ?? metabolism
2. Climate: ?? temperature = ?? metabolism
3. Physical activity: ?? activity = ?? metabolism
4. Genetics: hormone released from thyroid, located at bottom of neck

The new "metabolon" measures the hormones being released.

Discourse Analysis: Written Text

Student 1 received credit for remembering the underlined portions of the passage. Student 2 received credit for the same underlined information even though rice was not mentioned. Although other researchers might argue with this scoring decision, the important point is that such decisions need to be made on a principled and reliable basis. To achieve this, we have found it not only useful but necessary to keep a log of coding decisions so that inter- and intrarater reliability can be achieved.

In general, the task of deciding on making preserving is nontrivial because of the variability in how people recall information they can review, even when they are accurate. As these two student examples illustrate, people combine multiple sentences and produce summary statements; they combine ideas from parts of different sentences; they make accurate as well as inaccurate inferences; and they express what they remember in different structural forms. To the degree that learners combine information and add information that goes beyond what is present in the text, we would credit these recalls as reflecting situation-model-level understanding. (For a more detailed discussion, see Coté et al., 1998; Kintsch, 1994, 1998; McNamara & Kintsch, 1996; van Dijk & Kintsch, 1983; van Oostendorp & Goldman, 1999.)

In the analysis of the responses to "The Tupelo River Mystery" and "Metabolism" texts, we have attempted to model the kind of thinking that researchers need to do when they are attempting to understand what students have learned from reading a passage or set of passages on specific topics. You may have noticed that this thinking depends on a model of the conceptual domain as well as how that information is encapsulated in the specific passage. If one goal of reading informational text is to enable learners to acquire conceptual knowledge, the content and structure of that conceptual domain must be part of the discourse analysis of written text. This is as true for the analysis of passages and what they afford in the way of opportunities to learn concepts and relations in a particular domain as it is for understanding the situation models that are suggested by what learners write about these topics after reading.

Spontaneous Production of Written Discourse

The analysis of written text is often useful for assessing learners' prior knowledge, beliefs, feelings, and facility with the writing process. As indicated previously in our discussion, prior knowledge has a powerful impact on the way learners process, understand, and write about what they read. For informational text it is important to have a model of the conceptual domain in order to analyze both informational passages learners are provided with and what they learn as manifest in their written work. It is also possible to analyze essays that learners produce prior to any instruction in a topic area. These can be analyzed with respect to a conceptual model of the domain to estimate the contents and organization of learners' prestruc-
tion grasp of the domain. This information can be used in several ways, two of which are extremely important from the standpoint of literacy and instruction: as baseline assessments of domain understanding that can be compared with understanding at a later time and to select materials that are optimally suited to learners’ incoming knowledge level. (See Wolfe et al., 1998, for an alternative method for matching texts to readers.) Written productions can also be quite informative with respect to children’s cognitive and social development. Both content knowledge and knowledge of linguistic conventions, especially genre, are evident in children’s writing. For example, in a study by Chapman (1994, 1995), first graders’ productions during writer’s workshop were analyzed in terms of what genres they represented. Chapman noted that over the course of the year children’s writing became increasingly differentiated in terms of genre. Dyson (1997, 2003) examined the ways in which children construct their roles and relationships in text. In work with children in urban neighborhoods, she documented both challenges and adoption of gendered and cultural stereotypes (Dyson, 2003). From an instructional perspective, reader response logs are frequently advocated as a means by which children can develop engagement and enjoyment in text as well as comprehension and reflective processes (Barr, Blachowicz, Katz, Kaufman, & Wogman-Sadow, 2001).

**Subjective but Scientifically Sound**

Discourse analysis of written text is a necessarily subjective activity. Subjectivity does not, however, mean idiosyncratic or arbitrary. Scientifically sound discourse analysis follows principles and processes that ensure the rigor and reliability of the endeavor. For example, discourse analysis needs to follow the standard procedures for developing coding schemes that can be used reliably by multiple coders (e.g., Strauss, 1987; Strauss & Corbin, 1990). These procedures bring a greater degree of objectivity to the coding scheme. As mentioned previously, logs of decisions or difficult cases can help a rater maintain consistency across samples. In addition, it is standard practice to include some measure of interrater reliability. Reliability measures may be reported as simple correlations between raters, or when there is a finite set of possible observations and one can compute both hits and misses in the coding of observations, Cohen’s (1960) kappa may be used. The purpose of having a second person code the data is to ensure that codes are applied in a principled manner. If two people cannot concur in their use of a coding scheme, it seems more likely that the coding scheme represents an objective set of criteria for differentiating among examples, and that differences in coding are meaningful and not just arbitrary or subjective judgments. With such evidence of reliability, discourse analysis can fulfill its promise as a tool for description and distillation of important similarities and differences in written texts.

**AN EXEMPLAR OF THE ANALYSIS OF WRITTEN DISCOURSE**

In this section, we provide an exemplar study that has leveraged discourse analysis methodology in order to better understand how students process and learn from written text. Using this particular study (Wolfe & Goldman, 2005), we highlight four ways in which discourse analysis was used to support the exploration of how students learn from a set of two historical documents. Specifically, discourse analysis was used to:

1. Inform the design of the texts that students were given to read and to describe, characterize, and quantify.
2. Analyze performance on open-ended prior knowledge assessments.
3. Document-processing strategies during the reading of the texts.
4. Analyze the reasoning in student explanations generated after reading.

A final affordance was the ability to examine relationships among the measures, particularly through correlations of measures derived from the discourse analysis of comments made during reading the written texts and the measures derived from the discourse analysis of explanations that students generated after reading the texts.

In this exemplar, Wolfe and Goldman (2005) were interested in adolescents’ processing of information from multiple texts, reasoning about that information, and the relation between processing and reasoning. The texts represented two parallel accounts of a historical event, the Fall of Rome. Students first completed a prior knowledge assessment and then read the two texts, one after the other. Students thought out loud as they read each text. After reading the texts, students provided their own explanation for why Rome fell. The participants were sixth-grade students who had completed a 6- to 8-week unit on the Roman Empire approximately 2 months prior to the study. The materials consisted of two “historical documents” that were written by the researchers, a map of the Roman Empire (taken from the students’ textbook) with an accompanying time line of major events in the rise and fall of the Empire, and a list of facts that were “additional” to the information in the other materials.

**Design of Written Texts**

The first instance in which discourse analysis was used in this study was in the design of the two texts. These texts were written to appear as if they were historical interpretations developed by two different historians. Each provided a different perspective on why the Roman Empire could not defend itself during the Barbarian invasion of 476 A.D.: hedonism (instantiated in terms of laziness) or the sheer size of the empire and its implications
for defensibility. The texts were constructed to be structurally isomorphic, to agree on some information, including the Barbarian invasion, but to disagree as to the underlying cause of why the empire could not defend itself. Each text began with an explicit statement of disagreement with the theory presented in the other text. This was followed by a sentence indicating that the author had collected evidence to support his or her theory, followed by an explicit statement of his or her theory. Each text then developed a causal sequence of events that led to a conclusion: (1) the Roman Empire consisted of lazy citizens who could not effectively defend themselves, or (2) the Roman armies were too spread out to be able to defend against an attack on Rome. Each text cited a primary source in support of its argument (a painting or a diary). The last two sentences of each text stated that the Barbarians attacked and defeated the Roman Empire and reiterated the specific theory about the cause of the defeat. In addition to this attention to content and structure, the texts were also of comparable length and readability. The “lazy citizens” text was 12 sentences and contained 209 words, and had a Flesch-Kincaid grade-level score of 5.1. The “empire size” text was 15 sentences long, contained 203 words, and had a Flesch-Kincaid grade-level score of 5.5.

Open-Ended Prior Knowledge Assessments

Before reading the texts, students were asked what they remembered about the Roman Empire and why it was an interesting instance of discourse analysis in this study was to characterize and quantify these open-ended responses. This prior knowledge assessment was scored as the number of correct clauses provided about the Roman Empire. One point was given for each correct clause, and one-half point was given for the mention of a name or city without further information. Some example comments are as follows: “They copied their agriculture from the Greeks,” “If you are a woman you have very few rights,” and “They would watch chariot racing in the Coliseum.” Except for one student who was scored more than 3 standard deviations above the mean, most students were able to generate about four correct ideas about the Roman Empire.

Document-Processing Strategies

The third instance in which discourse analysis was used was in the coding of the think-aloud protocols as students read the written text. As a first step, the transcriptions of the think-alouds were divided into comments. A comment is the speech burst following the reading of a text sentence and corresponds to the coarse grain size used by Chi and colleagues (1994) in coding verbal protocols. Each comment was further parsed into events following procedures used by Coté and colleagues (1998) and corresponding to Chi and colleagues’ use of the idea unit. Events reflect different kinds of processing or ways of thinking about the text. For example, one student said, “I guess he disagrees with the other person because he thought that the army was too spread out. So he doesn’t think they covered too much territory.” This comment contains two events. The first event (first sentence) connects the content of this sentence to the other text and the other historian’s position. The second event (second sentence) is a paraphrase of the content of the sentence that was just read. Think-aloud data for all paragraphs were coded by two independent coders, with disagreements resolved through discussion. Interrater agreement on number of events was 90%.

Next, these think-aloud events were classified into one of five categories reflecting different kinds of processing, or thinking about the text: paraphrases (repeating gist), evaluations (positive or negative comments about the texts), comprehension problems (i.e., difficulty resolving pronoun references), comprehension successes (commenting when things make sense), and elaborations (adding meaning). These categories were chosen because prior research (Coté et al., 1998) but adapted to accommodate the multiple-text aspect of this task. Two independent raters agreed on the classification of 92% of the think-aloud events. Disagreements were again resolved in discussion.

Elaborations constituted the majority of the events (58%). Paraphrase events (22%) and evaluative events (13%) were also present. There were few comprehension-monitoring comments of either kind (3% each).

Because elaborations were the most frequent category of comments, two additional codes were developed in terms of (1) the source of the additional information and (2) the kind of processing involved. To determine the source of the information contained in elaborations, discourse analysis was used to identify overlap of information in comments with information from earlier in the text (same text); information from the other text, the map, or the time line (previous text); or information from a previous think-aloud comment. If there was no match with one of these sources, the information was attributed to prior knowledge. Almost half of the information used in elaborations was from prior knowledge (49%). However, the prior knowledge was overwhelmingly general world knowledge or personal experiences, not knowledge of the topic. The other sources of elaborations were distributed about equally across information from the same text (16%), the other text (18%), or a prior think-aloud comment (17%). What is significant in this pattern is that 34% of the elaborations involved connections between the focal sentence and information that was directly traceable to other text materials that were processed during the session, suggesting that students were attempting to compare and contrast across sentences.

The way that students made use of information in their elaborations was also coded. The three types of elaborative processing that encompassed almost all of the comments were self-explanation, surface text connection, and irrelevant associations. Elaborations were divided about equally
between associating (45%) and self-explaining (41%). Associative processing refers to connections that were not relevant to constructing a coherent understanding of the text (e.g., “My aunt visited Rome last year”). On the other hand, self-explanations involve meaningful integration of text content with prior knowledge or generation of a connection between text information that was not explicitly stated in the text. An example comment in this category is, “So they were afraid that something would happen because it was so far away or something.” This is a causal self-explanation because it adds a causal connection between a sentence from the text (“It was so far away”) and an emotional state being afraid, presumably based on prior knowledge of circumstances that generate fear. Thus, the elaboration used prior knowledge to make a causal inference. Only 12% of comments represented surface text connections in which the reader related one sentence to another sentence on the basis of surface features (“That talked about distance too”). Interrater agreement on coding of the elaboration events was 86%.

The relation between these two coding schemes was also informative because students used the different sources of information in different kinds of processing activities. This point is important because certain types of processing are more likely to contribute to the formation of coherent representations of the texts. For example, although students used prior knowledge in about half of the elaboration events, 70% of these were irrelevant associations, which may not contribute to understanding. Only 23% of the instances of prior knowledge use were in the context of causal self-explanations, a kind of processing that is more likely to contribute to building a coherent representation of the texts.

Explanations for the Fall of Rome

A fourth way that discourse analysis was used in this study was to code the open-ended explanations that students generated after reading. These explanations were coded along three dimensions: number of causes generated, complexity of reasoning, and integration of causes. The first coding dimension was a count of the number of causes included in the final explanation. These could be one or both of the causes from the texts (laziness of citizens or large size of empire) or causes taken from the additional information that was provided or generated from prior knowledge (e.g., “The Barbarians were too strong”). Beyond this simple count, the complexity dimension attempted to capture the variation in how students talked about the causes they listed. Some participants listed a cause or causes and wrote nothing else about them (these responses received no points for complexity). Other students provided additional information. Additional information that elaborated on the causes but did not explain why they were being mentioned received 1 point for complexity. Responses that explained why causes were being stated received 2 points for complexity. A final dimension of the coding focused on whether students related causes to each other. When participants’ responses integrated multiple causes by explaining the relation between (or among) the causes they had stated, they received an additional point. The points on these three dimensions were summed to give a score reflecting the level of reasoning demonstrated in these explanations. Ten of the explanations were coded by two raters, and agreement was 90% or better on each of the three dimensions. Differences were resolved through discussion. The remaining explanations were then coded by one of the raters. This analysis resulted in a mean reasoning score of 2.77, a mean number of causes of 1.48, and a mean complexity score of 1.18, indicating that students tended to do a little more than simply list the causes. However, only five students brought the alternative accounts together, explaining how the two causes from the texts combined to create the Fall of Rome.

Relations between Document-Processing Strategies and Explanation Coding

A final question of interest for this study was whether and how processing of the historical documents was related to the level of understanding that students achieved about the Fall of Rome. To address this question, the measures derived from the document-processing analyses were correlated with the reasoning scores. One main finding was that causal self-explanation during reading was related to better reasoning scores. A subsequent regression showed that causal self-explanations involving either prior knowledge or information from the other text were significant predictors of reasoning scores. Interestingly, noticing surface connections within the same text was also related to the reasoning score and contributed to the best fitting regression model. These surface connections may offer complementary mechanisms for constructing coherent representations. Furthermore, given the low correlation between prior knowledge self-explanations and surface connections, the data suggest that these alternative mechanisms are being used by different students. Finally, it is interesting to note that associative reasoning events (in which students made connections to irrelevant prior knowledge) were not predictive of reasoning scores. In other words, the discourse coding revealed specificity in the kinds of processing that benefited understanding of the topic.

Summary of Findings and Role of Discourse Analysis

The main goal of this exemplar study was to explore adolescents’ processing of information from multiple texts, reasoning about that information, and the relation between processing and reasoning. To achieve this goal, discourse analysis was pivotal in at least four areas of the methodology. First, it was used to create a set of reading materials with similarity on a number of critical dimensions. Second, it was used to evaluate open-ended
CONCLUDING COMMENTS

The approach to analysis that we have illustrated here provides a more complex picture of written discourse and students' understanding of it than is provided by a typical "right-answer" approach to scoring students' responses. Although more complex, the approach conveys information about specific concepts and ideas that students understand and how deeply they understand them. It also conveys information about how students' thinking differs from the information conveyed in the text, their misconceptions, and gaps in their understanding. As such, discourse-analytic approaches provide information that could guide instructional decision making in the classroom in ways that points on a multiple-choice or short-answer test cannot. In the future, we hope that the discourse-analytic approach will be more tractable for teachers to use in classroom instruction and assessment. Some research efforts are under way that are attempting to address this challenge through the use of various computational technologies (e.g., Foltz, Laham, & Landauer, 1999; Golden et al., 2003; Goldman, Golden, & van den Broek, 2007; Kintsch, Steinhardt, Stahl, & LSA Research Group, 2000; Lawless, Goldman, Gomez, Manning, & Braasch, in press; McNamar, Grasser, & Louworse, in press; Wolfe et al., 1998).

In closing, we wish to stress that researchers undertaking discourse analysis of written texts need to understand that it is an iterative process. It is driven by both theoretical and conceptual orientations and by the contents of the written discourse. In other words, even when researchers collect data with a priori ideas about the features or relationships that might be important, these ideas change as researchers engage in careful, scientific analysis of the written text that people produce. Discourse analysis is a process that successively approximates ways to adequately capture systematicity and variance across written texts and draw inferences about knowledge and learning.

REFERENCES


CHAPTER 8

Ethnographic Research

Victoria Purcell-Gates

Literacy researchers who operate out of a theoretical frame that views literacy as cultural practice are particularly drawn to ethnography as a methodological tool. This is because ethnography is grounded in theories of culture and allows researchers to view literacy development, instruction, learning, and practice as they occur naturally in sociocultural contexts.

Ethnography contributes to the quilt of research methodologies in that it allows literacy researchers to explore and come to understand phenomena about which little is known. As such, it provides those researchers who are so inclined a method for exploring and making sense of their data that is scientific and trustworthy. If done appropriately, the results, or outcomes, of ethnographies provide critical understandings of language and literacy in situ. They offer hitherto unknown maps and perspectives on literacy learning and development, without which teachers and researchers would be operating more or less blindly, in the dark, as they plan for and implement instructional strategies that "should" work according to other research paradigms.

Ethnographies provide the landscapes and the details of worlds. They aim to discover, understand, and describe human behavior holistically, as it occurs naturally within social and cultural contexts. In so doing, ethnographers can look for patterns and themes that ethnographic consumers can take away and use to enhance their own understandings of similar actors and contexts.

Although much, but not all, of the data collected by ethnographers is qualitative, ethnography is distinguishable within the category of qualitative research in that it is rooted in the concept of culture. This is not true of all qualitative research. As LeCompte and Schensul (1999b) emphasize, "Ethnography generates or builds theories of cultures—or explanations of