CHAPTER 6

The definite article *the* as a cue to map thematic information*

Morton Ann Gernsbacher and Rachel R. W. Robertson
University of Wisconsin-Madison

To comprehend a passage, readers and listeners (whom we refer to as "comprehenders") must build a mental representation of that passage. The Structure Building Framework (Gernsbacher, 1990; 1991; 1995; 1997) describes a few of the general cognitive processes and mechanisms that comprehenders use to build such a representation. According to the Structure Building Framework, the goal of comprehension is to build coherent mental "structures" and "substructures." These structures and substructures represent clauses, sentences, and passages. They are coherent in the sense that they share the main idea of the text, either locally or globally. This main idea is what is often called the topic, theme or "aboutness" of the text. Hence, the theme or topic of a passage is the main idea in the mental representation of that passage. It is therefore not surprising that the more coherent a passage is, the easier it is to build a coherent mental representation.

Comprehenders build each mental structure by first laying foundations (Carreiras, Gernsbacher, and Villa, 1995; Gernsbacher and Hargreaves, 1988; Gernsbacher and Hargreaves, 1992; Gernsbacher, Hargreaves, and Beeman, 1989). Then, comprehenders develop a mental structure by mapping incoming information onto the structure, when the incoming information coheres with the previous information (Deaton and Gernsbacher, in press; Gernsbacher, 1996; Gernsbacher and Givón, 1995; Gernsbacher and Robertson, 1992; Haenggi, Gernsbacher, and Bolliger, 1993; Haenggi, Kintsch, and Gernsbacher, 1995). However, if the incom-

ing information is less coherent, comprehenders shift and initiate a new substructure (Foertsch and Gernsbacher, 1994; Gernsbacher, 1985).

If we look at the construction of the theme in the text, we see the following: In terms of the Structure Building Framework, the theme of the text is initially formed in the processing stage of laying a foundation. It is in that stage comprehenders form hypotheses of what the text is about. In the next stage of the comprehension, comprehenders either map thematic information onto their foundational hypothesis, or they shift to lay the foundation for another hypothesis. We may assume that constructing a theme (building the hypothesis) will initially take more time than confirming it on the basis of incoming information. Furthermore, processing incoming information that confirms the hypothesis should be more rapid than processing information that requires a revision of the hypothesis or even a new hypothesis. Again, this can be fully explained by the Structure Building Framework: Laying a foundation will indeed consume additional mental effort. Furthermore, the process of mapping requires considerably less mental effort than the process of shifting.

Themes can be constructed at different levels of the text. The classic distinction is between local and global themes, the former operating at a clausal level, the latter at a discourse level. At the sentence level the "theme" is the sentence topic, often associated with subjectness or leftness (initial position) (Halliday, 1967). Gernsbacher and Hargreaves (1988) discovered such a processing advantage, what they called the Advantage of First Mention in sentences. At a more global level, themes can be found in paragraphs. Paragraphs are characterized by their thematic unity (cf. Hinds, 1977). The theme or topic of the paragraph can usually be found at the beginning of the paragraph (Hinds, 1980). Again, the effect of Advantage of First Mention can be found in clauses too, as has been shown by Gernsbacher, Hargreaves, and Beeman (1989). At larger sections like as episodes the same effect can be found. The reading times of the beginning sentence of an episode are considerably higher than those for other sentences (Haberlandt, 1984). In fact, Haberlandt's results demonstrate that average reading times for the first episode are longer than for the second episode. In short, there is evidence that the main idea of sections in the text, be it sentences, paragraphs or episodes, are expressed in the beginning of these sections. Clearly, this is what can be expected in terms of the Structure Building Framework. In laying a foundation additional mental effort is needed to build the theme of the text. The Advantage of First Mention predicts that the first part of a section is more accessible from comprehenders' mental representations. After the foundation has been laid, the comprehender maps thematic information onto the mental representation or shifts to a new thematic concept. The research we present in this chapter focused on the cognitive process we refer to as *mapping*.

According to the Structure Building Framework, comprehenders will map incoming information onto a mental structure when that incoming information coheres with the previous information. Mapping incoming information onto an existing structure or substructure takes less cognitive effort than shifting to initiate a new structure or substructure. Mapping incoming information onto an existing structure or substructure results in the incoming information being represented in the same structure or substructure as previously comprehended information. So, according to the Structure Building Framework, incoming information that coheres with previous information will be mapped onto the mental structure that represents the previous information. But what do we mean by coherence?

Dictionaries define coherence as "consistency," "continuity," and "co-ordination." Language researchers identify at least four sources of coherence in discourse: referential coherence, which is consistency in *who* or *what* is being discussed; temporal coherence, which is consistency in *when* the events that are being discussed occur; locational coherence, which is consistency in *where* these events occur; and causal coherence, which is consistency in *why* these events occur. The theme is expressed by the "who" and "what" of the text, rather than the "where" and "why", although the latter do support the former. This would mean that a psycholinguistic study of theme should focus on referential coherence. Referential coherence is the type of coherence we therefore explored in the experiments we present here. Two utterances can be considered referentially coherent and thus share thematic information if they refer to the same people, places or things.

So, one way to signal referential coherence is simply to repeat a word or phrase, for instance, repeating the words, *the authors,* in the following two sentences, *The authors were trying to illustrate their point. The authors were using an example.* In these two sentences, the repeated use of the words, *the authors,* suggests that the persons who were trying to illustrate their point were also the persons who were using an example. However, merely repeating a word does not ensure referential coherence; the word must refer to the same concept. For instance, these two sentences both contain the same words: *A reader was getting the point. A reader was getting bored.* However, it is unclear whether the reader who was getting the point was also the reader who was getting bored. If, however, the definite article *the* replaces the indefinite article *a* in the second sentence, *A reader was getting the point. The reader was getting bored,* this unfortunate situation is more apparent. Indeed, the definite article *the* can signal co-reference even when the noun it modifies is only a synonym of the previously mentioned noun, for instance, *A reader was getting bored. The member of the cultural elite put down the journal article and picked up a copy of* Harper's.

These examples illustrate how the English definite article *the* can signal referential coherence. Indeed, linguists and psycholinguists argue that the definite

article *the* indicates that the referent is "given:" the referent has been previously mentioned and its identity is known to the comprehender (Bock, 1977; Grieve, 1973; Haviland and Clark, 1974; Harris, 1974; Osgood, 1971).

For instance, Irwin, Bock, and Stanovich (1982: 308) write that "an important function of the article accompanying a referring expression is to indicate whether the expression has the same referent as information presented earlier in the discourse. In this capacity, the definite article *the* marks old, given, or presupposed information, while the indefinite article *a* marks new or asserted information."

Similarly, Murphy (1984: 489) writes that "a definite reference will in fact pick out . . . something that has been mentioned in the discourse or that is present in the [speakers' or writers' environmental] context (Clark and Marshall, 1981). Usually indefinite references introduce a new entity into the conversation. When the listener hears an indefinite article . . . , he or she can guess that a new entity is being mentioned."

In the research we present here, we investigated whether the definite article *the* acts as a cue to map thematic information onto the same mental structure. The starting point for our research was de Villiers' (1974) study. In de Villiers' (1974) experiments, two groups of subjects heard the same set of 17 sentences. For one group of subjects, the sentences occurred with only indefinite articles, for example: *A store contained a row of cages. A man bought a dog. A child wanted an animal. A father drove to his house. A cottage stood near a park.* For the other group of subjects, the same sentences occurred, but the indefinite articles were replaced with the definite article *the*, for example: *The man bought the dog. The child wanted the animal. The father drove to his house.*

When the sentences were presented with indefinite articles, subjects were more likely to interpret them as independent sentences that referred to multiple people and unconnected events. In contrast, when the sentences were presented with definite articles, subjects were more likely to interpret them as forming a coherent story in which the same persons and events were referred to repeatedly. These data by de Villiers (1974) suggest that the definite article *the* signals referential coherence.

The question we explored in our own research was whether this signal of referential coherence cues the structure building process of mapping in general and of mapping thematic information in particular. If so, then sentences like de Villiers (1974) presented should be read more rapidly when they contain the definite article *the* than when they contain indefinite articles. According to the Structure Building Framework, mapping incoming information onto a developing structure or substructure takes less cognitive effort than shifting to initiate a new structure or substructure. We tested this hypothesis in our first experiment.

Experiment 1

Our experimental stimuli comprised 10 different sets of sentences. Each set contained 14, 15, 16, or 17 sentences. We presented these sentences to two groups of subjects. One group of 24 subjects read all the sentences with indefinite articles, and the other group of 24 subjects read the sentences with the definite article *the*.

For example, one group of subjects read the following set of sentences:

(1) Some siblings were happy to be together.
(2) A road was icy and slick.
(3) A family stopped to rest.
(4) A cafe was almost deserted.
(5) A waitress took the order.
(6) A driver left to get gas.
(7) A man slipped and fell in a parking lot.
(8) A sister watched through a window.
(9) A gas station was nearby.
(10) An attendant rushed out of a building.
(11) A stranger helped a brother.
(12) A man walked slowly.
(13) A group stayed for a night.
(14) A trip was postponed.

The other group of subjects read the following set of sentences:

(1) The siblings were happy to be together
(2) The road was icy and slick.
(3) The family stopped to rest.
(4) The cafe was almost deserted.
(5) The waitress took the order.
(6) The driver left to get gas.
(7) The man slipped and fell in the parking lot.
(8) The sister watched through the window.
(9) The gas station was nearby.
(10) The attendant rushed out of the building.
(11) The stranger helped the brother.
(12) The man walked slowly.
(13) The group stayed for the night.
(14) The trip was postponed.

We measured how long the two groups of subjects spent reading each sentence. If the definite article *the* cues comprehenders to map, then the sentences should have been read more rapidly when they contained the definite article *the* than when

they contained indefinite articles. In addition, after subjects read each set of sentences they wrote down what they could remember from the sentences they just read. If the definite article *the* cues comprehenders to map, then the subjects who read the sentences with the definite article *the* should have been more likely to write sentences that suggested that the sentences had been integrated into one memory representation.

Method

Subjects. Forty-eight undergraduate students at the University of Oregon partici-pated to partially fulfill a course requirement. All subjects were native American English speakers. Twenty-four subjects were randomly assigned to the indefinite condition, and 24 were randomly assigned to the definite condition.

Materials. We wrote 10 different sets of sentences, modeled after the one set of sentences used by de Villiers (1974). Two sets contained 14 sentences; two sets contained 15 sentences, four sets contained 16 sentences; and two sets contained 17 sentences. The sentences ranged in length from four to eleven words.

Each set of sentences introduced two to three main characters who were subse-quently rementioned. Sometimes the main characters were rementioned with a verbatim repetition, for example, *The man slipped and fell in the parking lot. The man walked slowly.* Other times the main characters were rementioned with a different term, for example, *The driver left to get gas. The man slipped and fell in the parking lot.* or *The attendant rushed out of the building. The stranger helped the brother.* Each set of sentences also introduced one or two peripheral characters who were not rementioned, for example, *The waitress took the order.*

We created two versions of each set of sentences by making all the articles in one version indefinite *(a, an,* and, *some)*, and all the articles in the other version definite *(the).* Each of the two subject groups read only one version of the 10 sentence sets, either the version containing all indefinite articles or the version containing the definite article *the.* Both groups of subjects read the sentence sets in the same order.

Procedure. At the beginning of the experiment the subjects read instructions from a computer monitor. The instructions informed subjects that their task was to read several sets of sentences. They were told that the time they spent reading each sentence would be recorded but that they should read at their natural pace. The subjects were also told that after they read each set of sentences they would be required to write down as much as they could remember from the set of sentences

they just read. They were told that the more that they could remember the better and they should write down anything they could remember.

Before the first sentence of each set appeared, the subjects were warned by the word *READY?* , which appeared on the computer monitor. When the subjects pressed a response button to indicate that they were ready, the word *READY?* disappeared, and the first sentence of the set appeared. Each sentence was displayed in the center of the computer monitor. When subjects finished reading each sentence, they pressed a response button, and the next sentence of that set appeared.

At the end of each set of sentences, the following message appeared on the computer monitor: Please write in your packet as much as you can remember from this last set of sentences. When you are finished writing down as much as you can remember, press the response button to continue. Subjects were given a maximum of five minutes to recall as much as they could remember.

Each sentence (or partial sentence) that the subjects wrote was coded into one of eight categories: verbatim, nearly verbatim, synonym verbatim, referential verbatim, pronoun verbatim, partial, paraphrased, and integrative. An example of each category is given in Table 1.

Table 1. Example coding for subjects' recall of sentences in Experiment 1.

ORIGINAL SENTENCE: The man slipped and fell in the parking lot.

Verbatim: The man slipped and fell in the parking lot.

Nearly Verbatim: The man fell in the parking lot.

Synonym Verbatim: The man slid and fell in the parking lot.

Referential Verbatim: The driver slipped and fell in the parking lot.

Pronoun Verbatim: He slipped and fell in the parking lot.

Partial: The man slipped.

Paraphrased: The guy was walking and slipped on the ice.

ORIGINAL SENTENCES:
The driver left to get gas.
The man slipped and fell in the parking lot.
The attendant rushed out of the building.
The stranger helped the brother.

Integrative 2: The man was going to get gas, and on his way to the car he slipped and fell.

Integrative 3: The man was going to get gas, and on his way to the car he slipped and fell and the attendant rushed out.

Integrative 4: The man was going to get gas, and on his way to the car he slipped and fell and the attendant rushed out to help him.

Verbatim sentences were identical copies of the sentences that the subjects read (i.e., the subject wrote down all words of the original sentence using the exact wording). **Nearly Verbatim** sentences had the exact wording of the original sentence for all but one or two words; these one or two words could be additions or deletions, but not substitutions. **Synonym Verbatim** sentences also had the exact wording of the original sentence for all but one or two words, and the non-verbatim words were synonyms. However, the synonyms could not be reference terms (e.g., writing *the/a attendant* for *the/a stranger*). **Referential Verbatim** sentences had the exact wording of the original sentence for all but one or two words, and the non-verbatim words were reference terms. **Pronoun Verbatim** sentences had the exact wording of the original sentence for all but one or two words, and the non-verbatim words were pronouns. **Partial** sentences had missing or incorrect information, but at least one third of the words were verbatim. **Paraphrased** sentences did not have the exact wording of the original sentence, but the words conveyed the meaning of the original sentence. Finally, **Integrative** sentences captured the ideas of more than one of the original sentences. In addition, we identified how many original sentences were integrated into each integrative sentence. For example, an **Integrative 3** sentence captured the ideas of three of the original sentences. Two judges coded all of the subjects' responses, and the two judges agreed on 82 per cent of the coding; the remaining codes were assigned by consensus.

Results and Conclusions

First, we analyzed the subjects' reading time. Sentences were read more rapidly when they contained the definite article *the* (M=2419 ms; SE=23.7 ms) than when they contained indefinite articles (M=2960 ms; SE=26.1 ms), F_1 (1, 7478)=234; F_2 (1, 155)=333; $minF'$ (1, 873)=137.46. These data support the hypothesis that the definite article *the* cues the structure building process of mapping.

Second, we analyzed the subjects' recall performance. We found that subjects wrote the same number of sentences regardless of whether they had read the sentences with definite or indefinite articles. Both groups wrote, on the average, 8.8 sentences per sentence set. However, and more importantly, the types of sentences that the subjects recalled differed depending on whether they had read the sentence sets with definite article *the* or with indefinite articles.

Figure 1 presents the proportion of sentences of each type that subjects who read the sentences with the definite article *the* wrote (top circle) versus the proportion of sentences of each type that subjects who read the sentences with indefinite articles wrote (bottom circle). As Figure 1 illustrates, subjects who read the sentences with the definite article *the* did not differ from subjects who read the
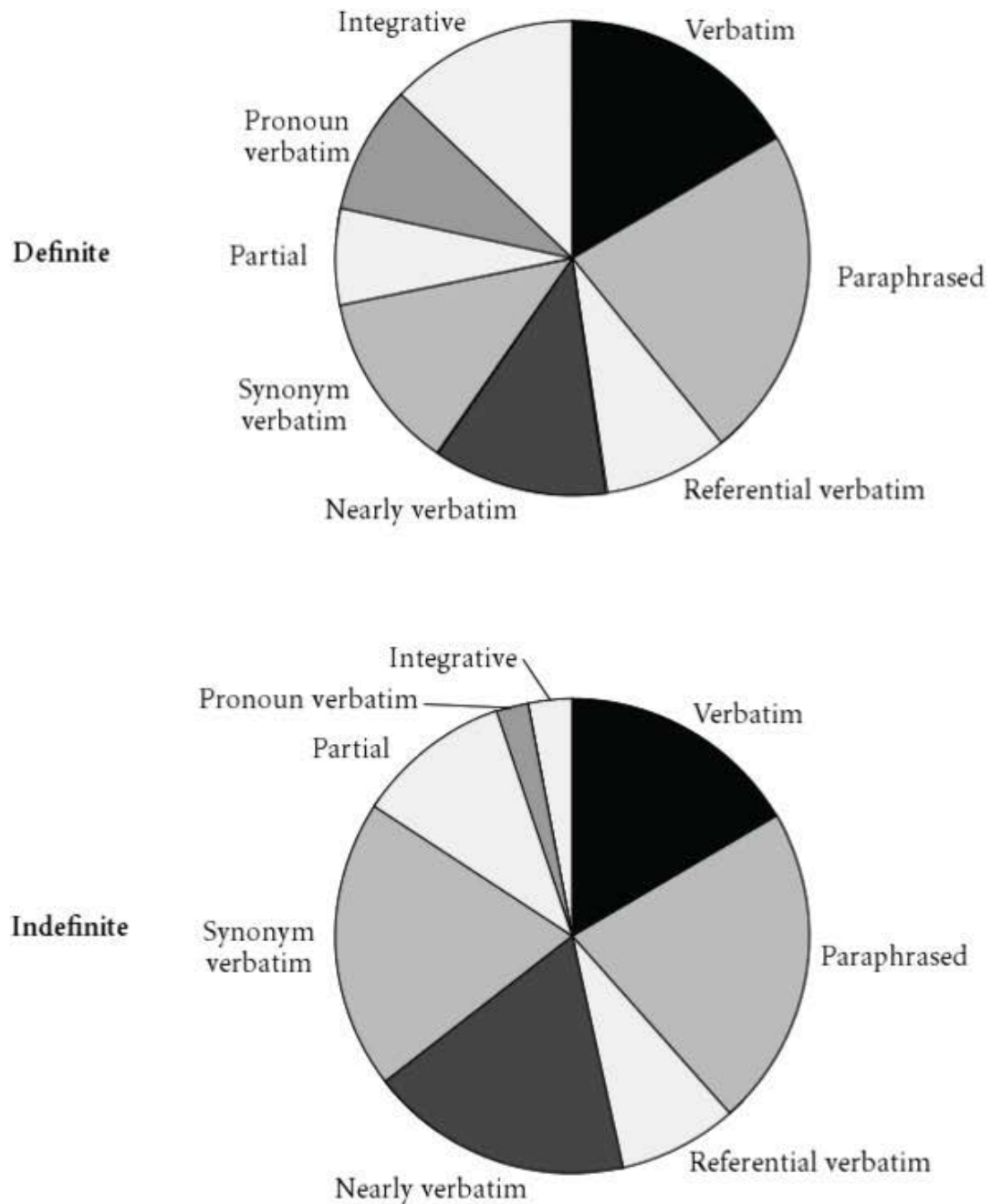
Figure 1. Proportion of sentences of each scoring type written by subjects who read the sentences with the definite article *the* (the top "pie") versus subjects who read the sentences with indefinite articles wrote (the bottom "pie"). Please see text for more information about the scoring categories

sentences with the indefinite articles in the proportion of **Verbatim, Paraphrased,** or **Referential Verbatim** sentences that they wrote (all $Fs < 1$). Subjects who read the sentences with indefinite articles wrote slightly more **Nearly Verbatim** sentences than did subjects who read the sentences with definite article *the*, although this difference was not reliable ($p > .09$). Subjects who read the sentences with indefinite articles did write reliably more **Synonym Verbatim** sentences, $F(1,$

477)=25.28, $p < .0001$. The greater production of Synonym Verbatim sentences by the subjects who read the sentences with indefinite articles was most likely because a sentence was considered a synonym verbatim sentence if the subject replaced an indefinite article with the definite article, and vice-versa. More frequently, subjects replaced an indefinite article with the definite article *the* (a pattern also reported by Luftig, 1981). In addition, subjects who read the sentences with indefinite articles wrote reliably more Partial sentences, $F(1, 477)=25.19$, $p < .0001$.

In contrast, subjects who read the sentences with the definite article *the* wrote significantly more Pronoun Verbatim sentences, $F(1, 477)=34.24$, $p < .0001$. And most strikingly, subjects who read the sentences with the definite article *the* wrote significantly more Integrative sentences, $F(1, 477)=52.42$, $p < .0001$, for Integrative 2 sentences; $F(1, 477)=15.80$, $p < .0001$, for Integrative 3 sentences; and $F(1, 477)=8.622$, $p < .0004$, for Integrative 4 sentences; and $F(1, 477)=84.57$, $p < .0001$, for all Integrative sentences. The difference between the number of Integrative sentences written by subjects who read the sentences with the definite article *the* versus the number of Integrative sentences written by subjects who read the sentences the indefinite articles was the largest difference we observed.

The finding that subjects who read the sentences with the definite article *the* wrote reliably more integrative sentences resembles a finding reported by Schultz and Kamil (1979). Prior to performing a recall version of the Bransford and Franks' (1971) "linguistic integration" task, half the subjects heard sentences that contained the definite article *the,* and half the subjects heard sentences that contained only indefinite articles, as typically occurs with the Bransford and Franks' stimuli. The subjects who heard the sentences with the definite article *the* were more likely to recall sentences that shared the same referent consecutively (even though these sentences were not presented consecutively in the acquisition list). The fact that subjects who read the sentences with the article "the" wrote significantly more integrative sentences also supports the idea the definite article supporting the construction of themes in the text. The main idea in a series of sentences is integrated in one sentence particularly when referential coherence can be established.

The higher incidence of referential "clustering" in Schultz and Kamil's (1979) subjects' recall and the higher incidence of integrative sentences in our subjects' recall support the hypothesis that comprehenders use the definite article *the* to map sentences onto the same mental structure. In our second experiment, we tested this hypothesis more directly.

Experiment 2

In Experiment 2, we again presented 10 sets of sentences to two groups of subjects. We again manipulated whether the articles in the sentences were indefinite or the definite article *the,* and we again measured subjects' reading times

for the sentences. However, in lieu of asking subjects to recall what they remembered after reading each set of sentences, we used McKoon and Ratcliff's (1980) priming-in-item verification task to measure how closely represented the sentences were in the subjects' mental structures.

More specifically, each time subjects read two sets of sentences, they performed a timed-verification task on a list of test sentences. Half the sentences in each test list were "true" sentences ("old" sentences that the subjects had read in one of the two recently read sets of sentences), and half sentences in each test list were "false" ("new" sentences that the subjects had not read in either of the two recently read sets of sentences). Unknown to the subjects, the test lists were constructed so that each "true" or "old" sentence was preceded in its test list by another "true/old" sentence. Furthermore, half the time, the preceding "true/old" sentence was from the same set of sentences as the following "true/old" sentence, and half the time the preceding "true/old" sentence was from the other set of sentences. In this way, we could measure how closely readers had mentally represented sentences in the same set of sentences compared with sentences in two different sets of sentences.

We predicted that subjects who read the sentences with the definite article *the* would be more likely to map the sentences of each set onto the same mental structure. If so, then subjects who read the sentences with the definite article *the* should have been faster to verify a "true/old" sentence when it was preceded by a sentence from the same set of sentences than when it was preceded by a sentence from a different set of sentences.

Method

Subjects. Seventy-two undergraduate students at the University of Oregon participated to partially fulfill a course requirement. Thirty-six subjects were randomly assigned to the indefinite condition, and 36 were randomly assigned to the definite condition.

Materials. The materials included the 10 sets of sentences we constructed for Experiment 1. Again, the sentences appeared in two versions: In one version all the articles were indefinite, and in the other version all the articles were the definite article *the*. Each of the two groups of subjects read only one version of the 10 sentence sets (the version with the indefinite articles or the version with the definite article *the*).

Each time that subjects finished reading two sets of sentences, they were presented with a list of test sentences. Because subjects read ten sets of sentences, they were,

therefore, presented with 5 lists of test sentences. Each list tested sentences from both sets of sentences that the subjects had just completed reading. Half the 32 test sentences in each list were "true," (sentences that the subjects had read before) and half were "false" (sentences that the subjects had not read before). The "false" sentences described characters and places that were described in the sentence sets, but the information conveyed in the "false" sentences was untrue. For example, in one set of sentences, subjects read that *The/A cafe was almost deserted.* A "false" test sentence for this set was *The/A cafe was crowded.* We wrote eight "false" sentences to match each sentence set.

The 16 "true/old" test sentences that appeared in each test list were copies of 16 sentences that subjects had read in the two sentence sets they just finished reading. Eight "true/old" test sentences were taken from one set, and eight were taken from the other set. Four "true/old" sentences per sentence set were target sentences, and four "true/old" sentences per sentence set were prime sentences. A "true/old" prime test sentence preceded each "true/old" target test sentence.

We created two versions of each test list by counterbalancing whether the prime sentence preceding each target sentence was from the same set of sentences or from the other (a "different") set of sentences. For example, the prime test sentence, *The/A stranger helped the brother* and the target test sentence, *The/A man walked slowly* were from the same sentence set. In contrast, the prime test sentence, *The/A student stood in line to board the/a plane* and the target test sentence, *The/A man walked slowly* were from two different sentence sets. In each test list, half the target sentences were primed by a sentence from the same set, and half were primed by a sentence from a different set. Across the two versions of the test lists, the prime sentences served both as primes for "same" target sentences and as primes for "different" target sentences. Thus, if a prime sentence in one version was from the same sentence set as the target sentence, in the other version it was from a different sentence set. The same test sentences occurred in both versions of the test lists, and the only difference between the two versions was the location of the prime sentences; in one test version a particular prime sentence occurred before a target sentence that was from the same set, and in the other test version that same prime sentence occurred before a target sentence that was from a different sentence set. In this way, the "true/old" target sentences and all the "false/new" test sentences occurred in the same position in the both versions of the test lists.

To summarize: Each of the 5 test lists comprised 32 test sentences. Sixteen test sentences were "false/new," and of these 16 "false/new" test sentences, 8 were based on each sentence set. The remaining 16 test sentences were "true/old," and of these 16 "true/old" test sentences, 4 "true/old" sentences were target sentences from one sentence set; 4 "true/old" sentences were target sentences from the other

sentence set; 4 "true/old" sentences were prime sentences from one sentence set; and 4 "true/old" sentences were prime sentences from the other sentence set.

All sentences in the test list matched the sentences that the subjects had read with regard to the articles (i.e., for subjects who read the sentences with indefinite articles, all test sentences appeared with indefinite articles, but for subjects who read the sentences with the definite article *the*, all test sentences appeared with the definite article *the*).

Procedure. As in Experiment 1, at the beginning of the experiment the subjects read instructions from a computer monitor. The instructions informed subjects that their task was to read several sets of sentences. They were told that the time they spent reading each sentence would be recorded but that they should read at their natural pace. The subjects were also told that after they read two sets of sentences they would be tested on how well they remembered those sentences. They were told that they would see a list of test sentences and for each test sentence they should decide whether that test sentence was true or false, based on the sentences they just read.

Before the first sentence of each set appeared, the subjects were warned by the words *READY FOR A SET OF SENTENCES?*, which appeared on their computer monitors. When the subjects pressed a response button to indicate that they were ready, the warning disappeared, and the first sentence of the set appeared. As in Experiment 1, subjects pressed a response button each time they finished reading a sentence, and the next sentence of the set would appear.

Each time subjects finished reading two sets of sentences, the words *READY FOR TEST SENTENCES?* appeared on the subjects' computer monitors. When the subjects pressed a response button to indicate they were ready, the warning disappeared, and the first test sentence appeared. Subjects responded to each test sentence by pressing a button labeled "TRUE," when they judged the test sentence to be true, or by pressing a button labeled "FALSE," when they judged the sentence to be false. After subjects responded to all the test sentences in a test series (i.e., 32 test sentences), the computer displayed each subject's percentage correct.

Results and Conclusions

First, we analyzed the subjects' reading times. Sentences were read more rapidly when they contained the definite article *the* (M=1973 ms; SE=11.5 ms) than when they contained indefinite articles (M=2085 ms; SE=12.0 ms), F_i (1, 11230)=55.6,
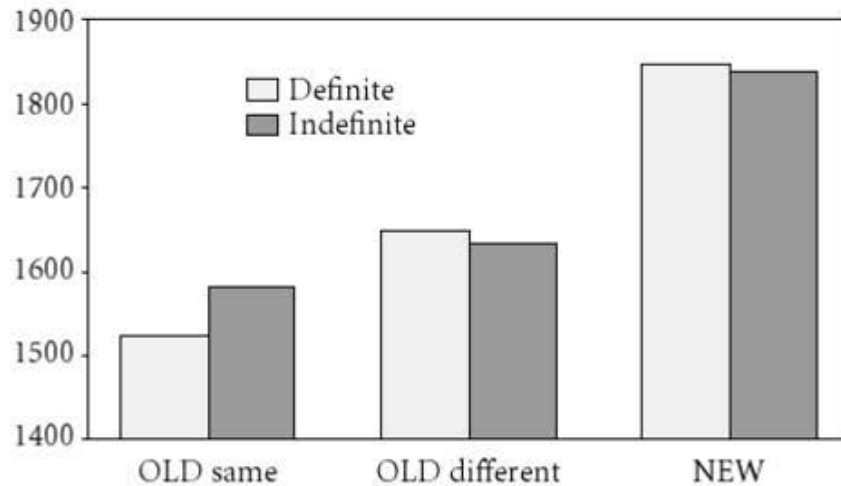
Figure 2. Average verification latencies to "true/old" sentences preceded by prime sentences from the same set of sentences, "true/old" sentences preceded by prime sentences from a different set of sentences, and "false/new" sentences

F_2 (1,155)=76.8, $minF'$(1,857)=32.26. These data support the hypothesis that the definite article *the* cues the structure building process of mapping.

Second, we analyzed the subjects' verification performance. Figure 2 displays the average verification latencies of subjects who read the sentences with the definite article *the* (represented by the unfilled bars) and subjects who read the sentences with indefinite articles (represented by the filled bars). As Figure 2 illustrates, subjects who read the sentences with the definite article *the*, verified "true/old" sentences more rapidly when those sentences were primed by sentences from the same set than when they were primed by sentences from a different set. This average 105 ms priming effect was statistically reliable, F_1 (1, 35)=16.52, F_2 (1, 39)=11.17, $minF'$(1,73)=6.66. In contrast, subjects who read the sentences with indefinite articles, did not verify "true/old" sentences more rapidly when the sentences were primed by sentences from the same set than when they were primed by sentences from a different set. This 34 ms average priming effect was not reliably different from what would be expected by chance, F_1 (1, 34)=2.97, F_2 (1, 39) < 1, $minF'$ < 1.

However, as Figure 2 also illustrates, both groups of subjects verified "true/old" sentences more rapidly than they rejected "false/new" sentences, $minF'$(1, 96)= 31.34, for subjects who read the sentences with the definite article *the*, and $minF'$(1, 98)=11.08, for subjects who read the sentences with the indefinite articles. Thus, both groups of subjects remembered the original sentences well enough to correctly reject sentences that they had not read. The aspect of performance in which the two groups differed was how much priming they received from sentences in the same versus a different set. Subjects who read the sentences with the definite article *the*

received a statistically reliable amount of priming from sentences from the same set; subjects who read the sentences with indefinite articles did not receive a statistically reliable amount of priming from sentences from the same set. These data support the hypothesis that subjects who read the sentences with the definite article *the* were more likely to map the sentences onto the same mental structure.

General discussion

In our first experiment, subjects who read sentences that contained the definite article *the* read those sentences significantly faster than did subjects who read the same sentences when they contained indefinite articles. When recalling the sentences that they had read, subjects who read sentences that contained the definite article *the* were also more likely to integrate several sentences into a single, composite sentence, and they were more likely to use pronouns instead of full noun phrases. In our second experiment, subjects who read sentences that contained the definite article *the* also read those sentences significantly more rapidly than did subjects who read the same sentences with indefinite articles. In addition, subjects who read the sentences with the definite article *the* verified the sentences that they had read more rapidly when those sentences were preceded by a test sentence from the same set of sentences rather than a test sentence from a different set of sentences.

Together, these results suggest that sentences that contain the definite article *the* are more likely to be mapped onto the same mental structure and should therefore be considered as highly relevant for thematic comprehension. According to the Structure Building Framework (Gernsbacher, 1990), comprehension involves building mental structures and substructures to represent discourse. When incoming information coheres with previously comprehended information, comprehenders map that information onto the structure or substructure that they are currently developing. Because the definite article *the* signals referential coherence, it cues comprehenders to map a mental representation of the sentence containing the definite article *the* onto the larger mental structure that represents previously read or heard sentences. This is what involves the establishment and development of a *theme* of the text.

Children as young as three adeptly interpret the definite article *the* as a signal of referential coherence. For example, three-year olds interpret the sequence *a doll* followed by *the doll* as referring to the same concept. Facility in producing the definite article *the* to convey referential coherence occurs just a bit later, around age four (Maratsos, 1976). The example of *a doll* followed by *the doll* illustrates a subtle shortcoming in our experimental materials. We chose to present two orthogonal conditions: one in which all the articles were indefinite and the other in which all

the articles were the definite *the;* however, our experimental sentences would have been more felicitous if we had reserved the definite article *the* for concepts' second mention. For example, instead of presenting the first sentence of one set of sentences as *The siblings were happy to be together,* and the third sentence of that set as *The family stopped to rest,* we could have presented the first sentence as *Some siblings were happy to be together,* and the third sentence as *The family stopped to rest.*

However, we predict that had we reserved the definite article *the* for only subsequent mention, we would have observed the same results in both of our experiments; indeed, the results might have been more striking, because we were following convention. Murphy (1984; Experiment 1) found that sentences containing the definite article *the* were read faster than the same sentences containing indefinite articles, even when the first mention of the co-referenced noun phrase was presented with an indefinite article. For example, after reading the sentence, *Though driving 55, Steve was passed by a truck,* subjects more rapidly read the sentence, *Later, George was passed by the truck, too* than they read the sentence, *Later, George was passed by a truck, too.*

Although our results, and those presented by Murphy (1984), might seem overly intuitive, a counter-hypothesis is that comprehending sentences that contain the definite article *the* should be more difficult than comprehending sentences that contain indefinite articles. Murphy (1984) states this counter-hypothesis in the following way: "An alternative hypothesis is that the definite article prompts the reader to search memory for the object being referred to, whereas the indefinite article indicates that a new object is being mentioned, and thus no such search is necessary . . . Therefore, indefinite reference should be easier to comprehend."

One problem with this counter-hypothesis is its assumption that comprehenders must search for the mental representation of a concept that is being co-referenced with the definite article *the.* We assert, instead, that speakers and writers use the definite article *the* to modify noun phrases whose concepts are most likely — rather than least likely — to be accessible in readers' and listeners' mental representations.

Another problem with this counter-hypothesis is its prediction that mapping a structure that represents a subsequent sentence onto a mental structure that represents a previous sentence is harder than shifting to initiate a new structure to represent that subsequent sentence. Our Structure Building Framework assumes just the opposite: Mapping coherent information onto a developing structure should be easier than shifting to initiate a new mental structure or substructure. In other words: holding on to an established theme should be easier than generating a new one. And a wealth of laboratory data supports this assumption (see Gernsbacher, 1990, Chapter 2).

A final problem with this counter-hypothesis is that it is unsupported by the data we presented here, as well as Murphy's (1984; Experiment 1) data. The experiments we reported here support the hypothesis that sentences containing noun phrases modified by the definite article *the* are more likely to be mapped onto the same mental structure. Because they are mapped onto the same mental structure, they are read more rapidly, they are recalled more integrally, and they are recognized more easily when they are cued by their neighboring sentences. Thus, the definite article *the* is used as a cue to map thematic information onto the same mental structure.

References

Bock, J.K. (1977). The effect of pragmatic presupposition on the syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior, 16*, 723–734.

Bransford, J.D., and Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331–350.

Carreiras, M., Gernsbacher, M.A., and Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin and Review, 2*, 124–129.

Clark, H.H. and Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B.L. Webber, and I.A. Sag (Eds), *Elements of discourse understanding*. Cambridge: Cambridge University Press.

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

de Villiers, P.A. (1974). Imagery and theme in recall of connected discourse. *Journal of Experimental Psychology, 103*, 263–268.

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

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

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

Gernsbacher, M.A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G.H. Bower (ed.), *The psychology of learning and motivation* (pp. 217–263). New York: Academic Press.

Gernsbacher, M.A. (1995). The Structure Building Framework: What it is, what it might also be, and why. In B.K. Britton, and A.C. Graesser (eds.), *Models of text understanding* (pp. 289–311). Hillsdale, NJ: Erlbaum.

Gernsbacher, M.A. (1996). Coherence cues mapping during comprehension. In J. Costermans, and M. Fayol (eds.), *Processing interclausal relationships in the production and comprehension of text* (pp. 3–21). Hillsdale, NJ: Erlbaum.

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

Gernsbacher, M.A., and Givón, T. (eds.). (1995). *Coherence in spontaneous text.* Philadelphia, PA: John Benjamins.

Gernsbacher, M.A., Goldsmith, H.H., and Robertson, R.R.W. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion, 6,* 89–111.

Gernsbacher, M.A., and Hargreaves, D. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27,* 699–717.

Gernsbacher, M.A., and Hargreaves, D. (1992). The privilege of primacy: Experimental data and cognitive explanations. In D.L. Payne (ed.), *Pragmatics of word order flexibility* (pp. 83–116). Philadelphia: John Benjamins.

Gernsbacher, M.A., Hargreaves, D., and Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language, 28,* 735–755.

Grieve, R. (1973). Definiteness in discourse. *Language and Speech, 16,* 365–372.

Haenggi, D., Gernsbacher, M.A., and Bolliger, C.M. (1993). Individual differences in situation-based inferencing during narrative text comprehension. In H. van Oostendorp, and R.A. Zwaan (eds.), *Naturalistic text comprehension: Vol. LIII. Advances in discourse processing* (pp. 79–96). Norwood, NJ: Ablex.

Haenggi, D., Kintsch, W., and Gernsbacher, M.A. (1995). Spatial situation models and text comprehension. *Discourse Processes, 19,* 173–199.

Harris, R.J. (1974). Effects of nonlinguistic knowledge on language production. *Journal of Verbal Learning and Verbal Behavior, 3,* 303–310.

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

Hinds, J. (1977). Paragraph structure and pronominalization. *Papers in Linguistics: International Journal of Human Communication, 10,* 77–99.

Hinds, J. (1980). Japanese expository prose. *Papers in Linguistics: International Journal of Human Communication, 13,* 117–158.

Irwin, D.E., Bock, J.K., and Stanovich, K.E. (1982). Effects of information structure cues on visual word processing. *Journal of Verbal Learning and Verbal Behavior, 21,* 307–325.

Luftig, R.C. (1981, August). *Normalization in paraphrase and recall.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

Maratsos, M.P. (1974). Preschool children's use of definite and indefinite articles. *Child Development, 45,* 446–455.

McKoon, G., and Ratcliff, R. (1980). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior, 19,* 369–386.

Murphy, G.L. (1984). Establishing and accessing referents in discourse. *Memory and Cognition, 12,* 489–497.

Osgood, G.M. (1971). Where do sentences come from? In D.D. Steinberg and L.A. Jakobovits (eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology* (pp. 497–529). Cambridge, UK: Cambridge University Press.

Schultz, E.E., and Kamil, M.L. (1979). The role of some definite references in linguistic integration. *Memory and Cognition, 7,* 42–49.