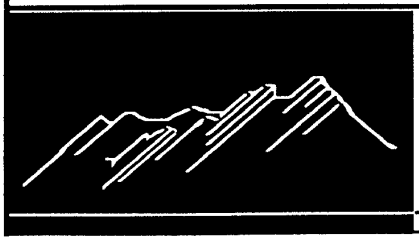
INSTITUTE OF COGNITIVE SCIENCE



Technical Report

University of Colorado, Boulder

Working Memory in Text Comprehension: Interrupting Difficult Text

Danielle S. McNamara & Walter Kintsch

Institute of Cognitive Science
University of Colorado
Boulder, Colorado 80309-0344

Technical Report #93-05

Working Memory in Text Comprehension: Interrupting Difficult Text

Danielle S. McNamara & Walter Kintsch

University of Colorado

Send Correspondence to:

Dr. Danielle S. McNamara
Institute of Cognitive Science, CB 344
University of Colorado
Boulder, CO 80309-0344
(303) 492-5236
mcnamara@clipr.colorado.edu

Abstract

We compare the effects of interrupting text dealing with familiar or unfamiliar domains with either arithmetic or sentence reading tasks. Readers were interrupted after each of the eight sentences, at the end of each sentence, or in the middle of each sentence. Previous findings of minimal effects of interruptive tasks on comprehension measures (e.g., Glanzer & Nolan, 1986) were replicated in this study. Also, as found by Glanzer and his colleagues, interruptions after each sentence of a familiar text by an unrelated sentence increased reading times by approximately 400 ms per sentence. In contrast, for difficult, unfamiliar texts, mid-sentence interruptions significantly lengthened reading times by 1262 ms for sentence and 1784 ms for arithmetic interruptions. These findings are explained in terms of Ericsson and Kintsch's (1993) memory model which proposes that skilled memory performance relies on the use of long-term memory as an extension of working memory, or long-term working memory.

Working Memory in Text Comprehension: Interrupting Difficult Text

Reading is by its very nature sequential. The glue of memory is needed, therefore, to hold the various elements of the sequence together. The eye moves from word to word, and in general each word is integrated with the previous ones as rapidly as possible (e.g., Just & Carpenter, 1987). According to Kintsch and van Dijk's (1978) model of text comprehension, sentences or phrases also form processing units, which are linked together via a short-term memory buffer. Evidence for the operation of such a buffer has been obtained in various experiments (e.g., Fletcher, 1981; Glanzer & Razel, 1974). However, reading comprehension is still possible when the use of this buffer is prevented. For instance, Glanzer and his colleagues (e.g., Fischer & Glanzer, 1986; Glanzer, Dorfman, & Kaplan, 1981; Glanzer, Fischer, & Dorfman, 1984; Glanzer & Nolan, 1986) have used an interruption procedure to interfere with short-term memory processes during reading. They did this by inserting unrelated material (e.g., unrelated sentences or arithmetic problems) after each sentence of a paragraph. The purpose of the intervening material was to interfere with the short-term retention of the just-read sentence. This interruption procedure produced far from dramatic effects. Although reading time for the next sentence in the paragraph was slowed by 300-400 ms, comprehension was totally unaffected.

Since the intervening sentence (or arithmetic problem) certainly must have interfered with integration processes in the short-term memory buffer, Glanzer's subjects must have found some way to continue reading without the use of that buffer. According to the theory of memory proposed by Ericsson and Kintsch (1993), this could be achieved through the use of *long-term working memory*. They assert that skilled memory performance relies on the use of long-term memory as an extension of working memory (i.e., consciousness or focus of attention). The theory of long-term working memory implies that all material in long-term memory that is connected via retrieval structures to cues available in working memory is directly accessible via a single retrieval operation, thus forming an extended, long-term working memory. Retrieval structures are generated during comprehension as an integral part of the comprehension process. Comprehension consists of forming mental representations (textbases and situation models in the theory

of van Dijk & Kintsch, 1983) which connect the various elements of the text representation in network-like structures. Thus, generating a text representation in itself creates a retrieval structure. Each successive sentence of a coherent text normally contains retrieval cues, such as related or repeated information (i.e., argument overlap), that provide access to that structure. Hence, the whole previous text structure is but a single retrieval operation away. Retrieval from long-term memory, if the retrieval cues are present in short-term memory, takes about 400 ms (e.g., Anderson, 1990; Yu et al., 1985). Glanzer's interruption procedure costs the reader, therefore, no more than a single retrieval operation, that is, about 400 ms.

A long-term working memory (Ericsson & Kintsch, 1993) explanation of the results obtained with the interruption procedure used by Glanzer (e.g., Glanzer & Nolan, 1986) implies that more serious disruptions of reading should be found (a) if there are no retrieval structures available, or (b) if there are no retrieval cues accessible in short-term memory. In the present experiment, an attempt was made to prevent, or at least interfere with, the formation of retrieval structures during comprehension. For this purpose, the interruption procedure used by Glanzer was elaborated in two ways. First, in one condition, sentences were interrupted in mid-sentence rather than at the end. Since the sentence is incomplete at this point, it is more likely that the partially constructed mental representation would not connect to the earlier portion of the text. Second, for some subjects, difficult texts from unfamiliar domains were used instead of easy, familiar texts. Since the construction of a situation model is strongly dependent on the availability of relevant background knowledge, the use of unfamiliar texts further decreases the likelihood that a workable retrieval structure can be generated. Neither one of these manipulations will completely prevent readers from some understanding - even difficult half-sentences will be understood to some degree. Nevertheless, a significant deterioration of understanding would be expected. To the degree that this happens: (a) no retrieval structure will be available; (b) the succeeding sentence fragment will not reinstate the previous text in long-term working memory with a single, 400 ms retrieval operation; and (c) more complex, time consuming, retrieval processes (e.g., deliberate search, construction of retrieval cues) will be required. If these retrieval processes are not successful, comprehension difficulties as well as longer reading times will result. If, on the other hand, these retrieval processes are successful, at the least the reading time for the second sentence half

will be lengthened by more than 400 ms. To obtain a more sensitive test of comprehension, a free recall test was used in the present experiment instead of comprehension questions, as in the original work by Glanzer and his collegues.

Method

Subjects and Design

The subjects were 72 undergraduate students at the University of Colorado who participated for course credit. A 2x2x3 mixed factorial design was employed, with two between-subjects factors, text domain (familiar, unfamiliar) and interference task (sentence, arithmetic), and one within-subjects factor, interruption type (control, end, middle). Eighteen subjects were assigned to each of the four between-subjects conditions.

Apparatus

Stimuli were presented either with Zenith Data Systems or IBM/PC computers. The Zenith computers were equipped with Zenith monitors, and the IBM/PC computers had Amdek 310 or 410 monitors. Subjects were also provided with pencil and paper to write down their recall of the text.

Materials

Text. The experimental texts consisted of 10 8-sentence paragraphs, as well as 100 unrelated sentences obtained from various sources such as encyclopedias, journals, and books. The paragraphs were chosen either in domains that were highly familiar to undergraduate students (e.g., "banana plants") or quite unfamiliar (e.g., "Fourier transformations") according to the judgment of the experimenters and two groups of norming subjects. Some of the paragraphs were reworded or altered to improve coherence or reduce length. The mean number of words per sentence for the familiar paragraphs, unfamiliar paragraphs, and unrelated sentences were 18.47, 16.12, and 17.02 words, respectively. Unrelated sentences were presented in random order, and each sentence was presented only once during an experimental session.

Arithmetic problems. The problems presented to the subjects in the arithmetic condition were one of three types: addition, subtraction, or multiplication. The type of problem presented to the

subject was randomly chosen. The addition problems were composed of two randomly chosen integers with the constraint that each integer was greater than 50 and less than 500. The two integers of the subtraction problems were randomly generated with the constraints that the first integer was greater than 550 and less than 1000, and that the second integer (i.e., the subtrahend) was greater than 50 and less than 500. The two integers of the multiplication problems were randomly generated with the constraints that neither integer was 10, the first integer was greater than 20 and less than 50, and the second integer was greater than 2 and less than 20. Thus, all answers to the three types of problems consisted of two or three digits and were less than 1000.

Interruption types. There were three interruption types: control, end, and middle. In the control condition, the paragraph sentences were presented in immediate succession, followed by a block of 10 interference tasks (i.e., either 10 unrelated sentences or 10 arithmetic problems). In the end condition, each sentence was followed by 1 interference task, except for the last sentence which was followed by the remaining 3 interference tasks. In the middle condition, each sentence was interrupted in the middle by an interference task; the remaining 2 interference tasks then followed the last sentence of the paragraph. An example of a text from an unfamiliar domain presented in the middle condition with unrelated sentences as the interference task is presented in Table 1. The example shows only the first four of the eight text sentences along with four of the ten interruption sentences. As shown in Table 1, the second half of an interrupted sentence began wherever it would have been located without the interruption.

Insert Table 1 about here

The interruption was placed between the subject and verb of the sentence, with the constraint that no fewer than three words (and at least two content words) preceded the break in the sentence. Sentences which began with "It is" were interrupted at a point such that there was an equal number of words in each of the two parts. If there were two equally important verbs in the sentence, the interruption was placed before the verb which resulted in the most equal division in the number of content words in the sentence.

Procedure

Subjects were tested individually in small rooms equipped with computers. All subjects read a series of 10 paragraphs, each consisting of 8 sentences. The first was a practice paragraph, presented in the middle condition. This text was identical for all subjects. A third of the remaining 9 paragraphs were presented in the middle condition, a third in the end-sentence condition, and a third in the continuous condition. Each of the nine texts was assigned to one of the three interruption types according to a Latin-square design. Thus, across subjects each text was presented in each interruption condition an equal number of times. The order of the texts presented to each subject was randomly generated.

Subjects were provided with oral instructions at the beginning of the session. They were told that written instructions would appear on the computer monitor and that they would complete a practice run to familiarize themselves with the experimental procedure. They were instructed to read the texts at their normal reading pace and not to attempt to memorize the sentences. They were also asked to read aloud the texts and the interference tasks (either the unrelated sentences or the arithmetic problems). Subjects were told to press the space bar after they had read each sentence aloud.

Subjects were told that interruptions would occur alternatively after the last sentence in the paragraph, between each sentence in the paragraph, and in the middle of each sentence. The subjects were told either that arithmetic problems or that unrelated sentences would appear as interruptions. They were informed that after they had read all of the sentences, they would be asked to write down as much of the text as they could remember. In the sentence interference task condition they were asked to recall both the sentences from the paragraph and the unrelated sentences.

The subjects in the arithmetic interference condition solved the arithmetic problems at their own pace and were allowed to use paper and pencil to solve the problems. As soon as the correct answer to the problem was entered, the next sentence of the text was presented. If they answered a problem incorrectly in less than 30 seconds, they were instructed on the computer monitor that their solution was incorrect and to try again. The same problem was then presented again. If, however, they answered a problem incorrectly after more than 30 seconds, they were informed that the answer was

incorrect, but were not asked to solve the problem again; they were presented with the next sentence of the text.

Results

Text Recall

For the purpose of scoring recall protocols each text was divided into idea units. For each unit subjects were given 1 point if they recalled the main gist of the unit and half of a point if they recalled only a fraction of the idea unit.

An analysis of variance was performed on proportion recall including the two between-subjects factors, text domain (familiar, unfamiliar), and interference task (sentence, arithmetic), and one within-subjects factor, interruption type (control, end, middle). There was a main effect of the familiarity of the text, $F(1,68) = 61.2$, $p < .001$, reflecting greater recall for the familiar text ($\underline{M} = 0.30$) than for the unfamiliar text ($\underline{M} = 0.16$). There was also a main effect of interference task, $F(1,68) = 42.0$, $p < .001$, reflecting greater recall when the texts were interrupted by the math problems ($\underline{M} = 0.29$), than by the sentences ($\underline{M} = 0.18$). The interaction between text and interference task was not statistically significant, $F(1,68) = 2.1$, $p < .157$. There was no effect of interruption type, $F(2,67) < 1$, nor did interruption type interact with either of the between-subjects variables. Thus, subjects recalled the same amount of text regardless of whether the interference tasks occurred at the end of the paragraph ($\underline{M} = 0.23$), at the end of each sentence ($\underline{M} = 0.24$), or in the middle of each sentence ($\underline{M} = 0.24$).

Sentence reading time

Sentence reading times by text domain (familiar, unfamiliar), interference task (sentence, arithmetic), interruption type (control, end, middle) are presented in Table 2.

Insert Table 2 about here

An analysis of variance was performed on sentence reading times including the two between-subjects factors, text domain (familiar, unfamiliar), and interference task (sentence, arithmetic),

and two within-subjects factors, interruption type (control, end, middle), and sentence order (1-8). Neither text domain, $F(1,68) < 2$, nor interference task, $F(1,68) = 2.4$, $p < .126$, nor the interaction of the two, $F(1,68) < 1$, had reliable effects on sentence reading times.

There was, however, a significant main effect of interruption type, $F(2,67) = 21.2$, $p < .001$. This finding reflects both longer sentence reading times for sentences interrupted in the middle ($\underline{M}$ = 10.72 s), compared to both the end and control interruption conditions ($\underline{M}$ = 9.85), $F(1,68) = 25.6$, $p < .001$, and longer sentence reading times for those interrupted at the end of each sentence ($\underline{M}$ = 10.17) compared to those which were not interrupted at all ($\underline{M}$ = 9.54), $F(1,68) = 12.0$, $p < .001$. Thus, the average interruption effect (i.e., the overall increase in reading time compared to the control condition) was 1180 ms when the sentences were interrupted in the middle, versus 630 ms when the sentences were interrupted at the end. Neither text domain, $F(2,67) = 2.5$, $p < .095$, nor interference task, $F(2,67) < 2$, reliably interacted with interruption type. However, a planned-comparison test showed that the contrast between the middle interruption and the two other interruption types (i.e., control and end) interacted significantly with text, $F(1,67) = 5.0$, $p < .030$, reflecting the finding that the middle interruption had a greater effect on reading time for the unfamiliar texts than for the familiar texts (see Figure 1). The contrast orthogonal to this comparison, between the end and control interruption conditions, did not reliably interact with text familiarity, $F(1,68) < 1$.

Insert Figure 1 about here

There was also a significant main effect of sentence order, $F(7,62) = 34.8$, $p < .001$, reflecting greater sentence reading times for the sentences at the beginning than at the end of the paragraph. Sentence order did not reliably interact with interference task, $F(7,62) < 2$, nor with interruption type, $F(14,55) < 2$, but did significantly interact with text familiarity, $F(7,62) = 12.8$, $p < .001$. As can be seen in Figure 2, which presents sentence reading times for unfamiliar and familiar texts, it takes longer to read sentences at the beginning of a text dealing with an unfamiliar domain than with an easier, more familiar text domain, whereas these differences diminished towards the end of the paragraph.

Insert Figure 2 about here

For the condition in which sentences were interrupted in the middle, we were interested in whether there were differences in reading time for the two parts of the sentence. The beginning of the sentence was uninterrupted, and thus should take less time to read, than the second, interrupted part of the sentence. We were also interested in whether this difference was augmented by the familiarity of the text or by the type of interference task. A separate analysis of variance was performed for only the middle interruption condition including two between-subjects factors, text domain and interference task, and one within-subjects factor, sentence part (interrupted, uninterrupted). There was neither a main effect of text, $F(1,68) = 2.7$, $p < .110$, nor an effect of interference task, $F(1,68) = 2.7$, $p < .106$. As predicted, there was an effect of sentence part, $F(1,68) = 89.8$, $p < .001$, reflecting shorter reading times for the beginning of the sentence (M = 4770 ms) than for the last interrupted part of the sentence (M = 5949 ms). This difference did not depend on the familiarity of the text, $F(1,68) < 1$, but did interact with the type of interference task, $F(1,68) = 12.7$, $p < .001$. It took longer to read the second part of the sentence following an arithmetic task (M = 6494 ms) than after reading an unrelated sentence (M = 5404 ms).

Interference tasks

An analysis of variance was performed on task completion times including the two between-subjects factors, text domain, and interference task, and two within-subjects factors, interruption type, and task order (1-10). The mathematics problems required significantly more time to complete (M = 14.8 s) than the sentences (M = 8.4 s), $F(1,68) = 153.5$, $p < .001$; and there was an effect of the order of the tasks, $F(9,60) = 4.2$, $p < .001$, reflecting a greater amount of time spent on the first task (M = 12.9 sec) compared to the remaining nine tasks (M = 11.5). No other effects or interactions were significant (all Fs < 2).

Discussion

The main findings of Glanzer and his collegues (e.g., Glanzer & Nolan, 1986) were replicated in the present study: Interruptions by an unrelated sentence following each sentence of a familiar paragraph had no effect on comprehension (here measured by a free-recall test), but increased reading times by 411 ms per sentence. Placing the interruptions in the middle of each sentence did not greatly change the results, as long as the reading material was familiar. However, for unfamiliar texts, mid-sentence interruptions significantly lengthened reading times by 1262 ms for sentence and 1784 ms for arithmetic interruptions, without affecting free recall.

This pattern of results remained the same when the texts were interrupted by arithmetic operations instead of unrelated sentences, except that considerably larger increases in reading times were obtained[1]. It may be the case that switching from arithmetic to reading requires a constant switching time which is responsible for this increase. Recall was much better in the arithmetic condition than in the sentence condition, in spite of the fact that the arithmetic tasks actually took longer on the average than reading the interpolated sentences. This result has to be expected, in part because reading unrelated sentences produces more verbal interference, and in part because in the sentence task subjects had to recall both the main paragraph and the interpolated sentences.

The present study thus confirms previous results obtained with the interruption procedure of Glanzer. It also extends these results by testing a critical prediction of the long-term working memory theory of Ericsson and Kintsch (1993) against the alternative interpretation of these data that has been offered by Glanzer (e.g., Glanzer & Nolan, 1986). The reading interruption procedure used by Glanzer does not impair comprehension and merely results in a relatively modest increase in reading time. Glanzer explained this finding by assuming that readers have access to a verbatim memory trace (and not thematic information), even after an interruption, that allows them to resume normal processing. This interpretation does not account for the significant interaction between text familiarity and mid-sentence interruption: If what is reinstated after the interruption were a raw, uninterpreted, verbatim trace of the sentence, this trace would be equally available for familiar and unfamiliar texts. On the other hand, the theory of long-term working memory predicts just such an interaction. According to models of discourse comprehension (e.g., van Dijk and Kintsch, 1983; Kintsch,

1988), as well as other structure-building models (e.g., Gernsbacher, 1990), comprehending a text involves the construction of a coherent mental text representation allowing the text to be accessed in long-term working memory via a single retrieval operation. This retrieval operation takes about 400 ms if the appropriate retrieval cues are present in working memory. Ericsson and Kintsch (1993) have reviewed data indicating that retrieval from long-term memory requires 1 to 2 seconds when the appropriate cues are not in the focus of attention (e.g., Charness, 1976; Ericsson & Staszewsky, 1989). In the present experiment, the mid-sentence interruption effect for unfamiliar texts was 1.3 seconds in the sentence condition and 1.8 seconds in the arithmetic condition. Thus, this delay is about what one would expect for long-term memory retrieval - in contrast to the 400 ms observed for retrieval using long-term working memory. If readers are interrupted in mid-sentence when they are reading an unfamiliar text for which they lack the knowledge to access readily the information needed to construct a situation model, they cannot generate a coherent text representation and hence do not have a retrieval structure to reinstate the previous text after the interruption. They must, therefore, use strategic retrieval operations. One example of such a retrieval operation would be a deliberate search for background knowledge (i.e., situation knowledge) and/or previous sentence fragments to be integrated with the new sentence fragment. Skilled readers are generally quite capable of adopting such strategies, but these mental operations are much more time consuming than the 400 ms retrievals involving long-term working memory.

Interrupting reading with an unrelated sentence or arithmetic problem must interfere with the short-term memory buffer. Nevertheless, as long as readers are able to understand a text (i.e., form a coherent mental representation of the text), interrupting reading merely lengthens reading time for the next sentence by the amount required for a single long-term working memory retrieval operation. This is true even when subjects read unfamiliar texts, or when the interruption occurs in mid-sentence rather than at the end of each sentence. In either case, apparently, readers are still capable of forming mental representations of the sentence (or sentence fragment) that can serve as efficient retrieval structures. Only when readers were given unfamiliar texts combined with mid-sentence interruptions was their comprehension impeded to such an extent that the preconditions for retrieval from long-term working memory were no longer present. In this case, subjects had to rely upon

strategic retrieval operations to access information from long-term memory, requiring significantly more time.

References

Anderson, J. R. (1990). Cognitive psychology and its implications. San Francisco: Freeman.

Charness, N. (1976). Memory for chess positions: Resistance to interference. Journal of Experimental Psychology: Human Learning and Memory, 2, 641-653.

Ericsson, K. A., & Kintsch, W. (1993). Memory in comprehension and problem solving: A long-term working memory. To be submitted for publication.

Ericsson, K. A., & Staszewski, J. (1989). Skilled memory and expertise: Mechanisms of exceptional performance. In D. Klahr, & K. Kotovsky (Ed.), Complex information processing: The impact of Herbert A. Simon (pp. 235-268). Hillsdale, N. J.: Erlbaum.

Fischer, B., & Glanzer, M. (1986). Short-term storage and the processing of cohesion during reading. Quarterly Journal of Experimental Psychology, 38, 431-460.

Fletcher, C. (1981). Short-term memory processes in text comprehension. Journal of Verbal Learning and Verbal Behavior, 20, 264-274.

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

Glanzer, M., Dorfman, D., & Kaplan, B. (1981). Short-term storage in the processing of text. Journal of Verbal Learning and Verbal Behavior, 20, 656-670.

Glanzer, M., Fischer, B., & Dorfman, D. (1984). Short-term storage in reading. Journal of Verbal Learning and Verbal Behavior, 23, 467-486.

Glanzer, M., & Nolan, S. D. (1986). Memory mechanisms in text comprehension. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 275-317). New York: Academic Press.

Glanzer, M., & Razel, M. (1974). The size of the unit in short-term storage. Journal of Verbal Learning and Verbal Behavior, 13, 114-131.

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

Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. Psychological Review, 85, 363-394.

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

Yu, B., Zhang, W., Jing, Q., Peng, R., Zhang, G., & Simon, H.A. (1985). STM capacity for Chinese and English language material. Memory & Cognition, 13, 202-207.

Author Notes

We are grateful to Anders Ericsson for many helpful discussions and to Eileen Kintsch and Timothy Rickard for their comments on this manuscript. We acknowledge and thank Ernest Mross for his assistance in the computer programming of this experiment and Julie McAfee for her help in conducting the experiment.

Footnotes

[1] Glanzer et al. (1981) did not find a corresponding increase in reading time, perhaps because their addition problems were much simpler (e.g., 31 + 8) and were not read aloud.

Table 1

Text example from an unfamiliar domain presented in the middle interruption condition with unrelated sentence interference tasks.

Metamorphic rocks are those that remain in the solid state

The probability of three events must be less than the probability of one of them alone.

while being changed

by heat and/or pressure, with or without overall chemical change.

Most metamorphic rocks crystallize under stress,

The odd thing about these deep and personal connections among women is that they often leap barriers of age, economics, worldly experience, race, and culture.

resulting in characteristic

foliation or parallelism of the constituent grains, especially micas.

The contact metamorphic rocks, however, more commonly

For parents who wish to start investing on a child's behalf, few vehicles could be more appropriate than a mutual fund.

form without

deformation, resulting in a massive texture.

Their proximity to a heat source and their characteristic spotted appearance resulting from the growth of new minerals

Thus the period became a somber and reflective one, looking forward toward the coming suffering of Christ.

aid in their recognition.

Note. The example shows only the first four of the eight text sentences along with four of the ten unrelated sentence interference tasks. Every sentence (or sentence half) appeared on the screen one at a time. The second half of an interrupted sentence began wherever it would have occurred on the screen without the interruption.

Table 2

Average sentence reading times in sec by text (familiar, unfamiliar), interference task (sentence, arithmetic), and interruption condition (control, end, middle)

	Familiar		Unfamiliar	
	Sentence	Arithmetic	Sentence	Arithmetic
Control	8.700	9.785	9.642	10.028
End	9.110	10.785	9.891	10.887
Middle	9.243	10.917	10.904	11.812

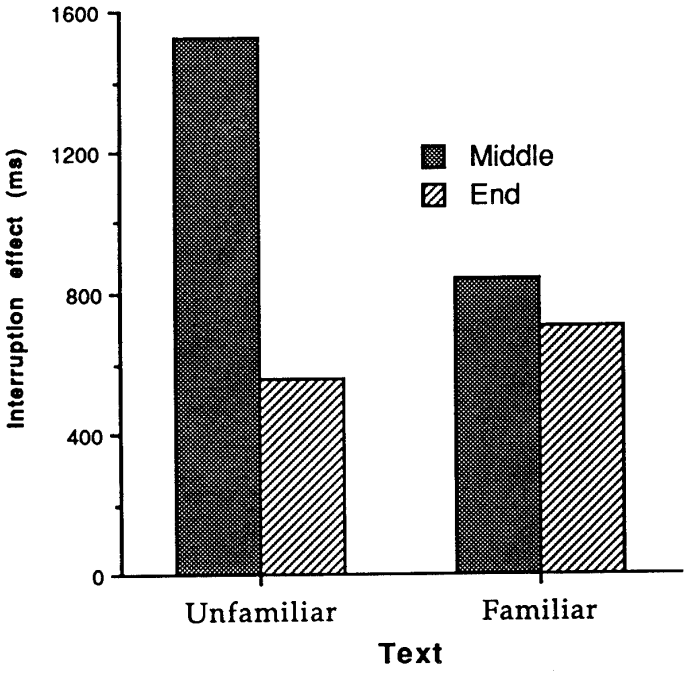


Figure 1. Interruption effects in reading time (i.e., difference from control reading times) for the middle and end interruption conditions by text (averaged over the arithmetic and sentence interference task conditions).

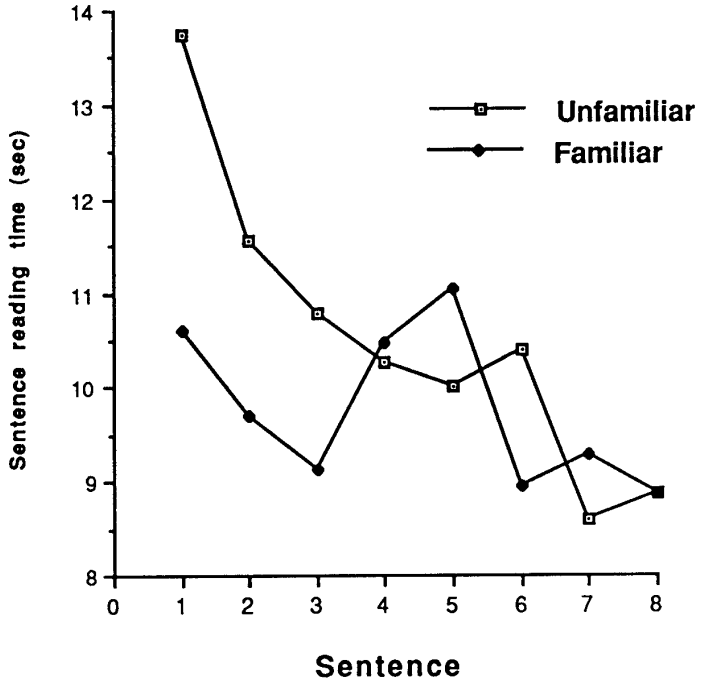


Figure 2. Sentence reading times by text.

---

[1] Glanzer et al. (1981) did not find a corresponding increase in reading time, perhaps because their addition problems were much simpler (e.g., 31 + 8) and were not read aloud. We are not able to examine any effects of problem difficulty in our study because the problems themselves were not recorded by the computer.