JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 22, 121–132 (1983)

The Activation of Antecedent Information during the Processing of Anaphoric Reference in Reading

GARY S. DELL

Dartmouth College and Max-Planck-Institut für Psycholinguistik

AND

GAIL MCKOON AND ROGER RATCLIFF

Yale University

The time course of the retrieval of antecedent information during the processing of anaphoric reference was examined in four experiments. In each experiment subjects read paragraphs that appeared one word at a time on a CRT screen. At unexpected times, they were given a single-word recognition test. Response times and error rates for these tests indicated that both the referent (e.g., *car*) of a superordinate anaphor (e.g., *the vehicle*) and concepts in the same proposition as the referent become activated as early as 250 milliseconds after the anaphor is read. The referent remains activated as the sentence is read, but the activation of other concepts dies away. The results are interpreted as support for the proposal that antecedent information is initially retrieved in the form of propositions, but only certain concepts from these propositions, those that are important for establishing text structure, remain activated.

One feature of a coherent text is that many of its sentences refer to concepts introduced earlier in the text. A concept such as *a car* might later be referred to as *the car, the vehicle,* or simply *it*. This type of reference is called *anaphoric,* or backward, reference.

Anaphoric reference plays a critical role in propositionally-based theories of text comprehension, which define the structure of a text as an interconnected set of propositions (e.g., Anderson, 1976; Anderson & Bower, 1973; Kintsch, 1974; Rumelhart, Lindsay, & Norman, 1972). According to these theories, two propositions are directly connected if they have at least one concept in common, that is, if the two propositions are statements about the same thing. Thus, if a comprehender, whether a listener or a reader, is to recover the text structure, he or she must be able to recognize when concepts are repeated. Upon reading an anaphor (a word that refers to a previous concept), the comprehender must identify its referent and establish the connection between the propositions containing the anaphor and the referent.

Our experiments address two issues concerning the processing of anaphoric reference in reading. First, exactly what gets retrieved from memory when the anaphor is first encountered? Second, what is the time course of this retrieval process?

One way to restate these issues is in terms of activation. When a reader encounters an anaphor, what are the concepts that then become activated, and when, and for how long, are they activated? Presumably, at some point after the anaphor is read, its referent becomes activated, reflecting successful retrieval. Both McKoon and Ratcliff (1980) and Chang (1980) have shown that an anaphor activates its referent at least by the end of the sentence contain-

This research was supported by NICHD Grant HD-13318-01. Reprint requests should be sent to Gary Dell, Department of Psychology, University of Rochester, Rochester, N.Y. 14627.

0022-5371/83/010121-12$03.00/0

ing the anaphor. In McKoon and Ratcliff's study subjects read short paragraphs. The last sentence of each paragraph contained either a superordinate anaphor (e.g., *the vehicle*) which referred to a word (e.g., *car*) in the first sentence, or contained no anaphor. A recognition test for the referent (*car*) was given immediately following the final sentence. It was found that recognition responses were faster when the final sentence contained the anaphor than when it did not, and it was concluded that the anaphor activates its referent. Chang's (1980) experiments and conclusions were similar. He found that recognition responses for a person's name (e.g., *John*) were faster when a final clause contained a pronoun referring to the person than when it did not.

Is the referent the only concept that is activated by processing an anaphor? Kintsch and Vipond (1978) proposed that when a concept from a text is activated, both the concept and the propositions containing it are activated. In other words, information about concepts is really retrieved in the form of propositions, rather than just concepts themselves. Applying this idea to the retrieval of referents implies that an anaphor should activate the proposition containing the referent as well as the referent itself. In particular, one would expect that concepts in the same proposition as the referent (hereafter called the referent's *companions*) would be activated by an anaphor. As an example, consider the following two sentences (from McKoon & Ratcliff, 1980):

A car sped around the corner out of control.
The vehicle smashed into some empty boxes.

The anaphor *vehicle* should activate the referent *car* and its companions *corner* and *sped around*, if the entire proposition is retrieved. McKoon and Ratcliff (1980) found that this was indeed the case. When a companion to the referent was given as a test word following the sentence mentioning the anaphor, response times were speeded relative to the condition in which there was no anaphor in the final sentence. Thus there is evidence that the entire proposition is activated.

There are, however, many unanswered questions regarding the retrieval of antecedent information during text processing. Specifically one would like to know the time course of the retrieval process. McKoon and Ratcliff's (1980) and Chang's (1980) experiments tested words after the clause containing the anaphor was read. Does the activation of referents wait until the end of the clause as might be expected from the notion that most significant grammatical processing occurs at clause boundaries (e.g., Bever & Hurtig, 1975)? Or does it occur immediately when the anaphor is encountered? Also, do the referent and its companions become activated at about the same time? One would expect so if it is truly the proposition that is retrieved.

We attempted to answer these questions by adapting the item recognition paradigm used by McKoon and Ratcliff (1980) so that we could give an unexpected single-word recognition test at any time while the subject reads the paragraph. McKoon and Ratcliff's subjects read paragraphs that were presented one sentence at a time on a CRT screen and the subject controlled his or her reading speed. The test word appeared after the last sentence was read. In order to be able to test recognition in the middle of sentences, we present words one at a time, with each word in a sentence turning on 250 milliseconds after the previous one. The whole effect is something like a neon sign, which starts out blank, and then slowly reveals its message a word at a time. By the end of the sentence all the words are displayed. Then the sentence disappears and the next sentence is presented in the same way. In this way, a test word can be presented at any point, after any word of the text. A test word is presented all in capital letters and underlined by asterisks, and when it is presented any part of the sentence that is on the screen disappears. With this procedure, it is possible to measure the state of activation of a previously presented word at a specific

point by presenting that word for a recognition decision.

The four experiments reported here used the paragraphs from McKoon and Ratcliff (1980), one of which is given in Table 1. Each paragraph occurred in a version which contained a superordinate anaphor in the last sentence referring to a word from the first sentence, and in a version which had no anaphor in the last sentence. Across the experiments we tested the activation of the referent (*burglar*), a companion (*garage*), and a control word from a different sentence (*bottles*), with both versions of the paragraph, and at various points in the final sentence. These test positions are labeled by numbers in the example in Table 1. Essentially we were interested in the speedup or "priming" in recognition responses brought about when the final sentence mentioned the anaphor (*the criminal*) over that when the final sentence did not mention an anaphor (*a cat*). To the extent that there is priming at a given test position with a given test word (referent, companion, or control), we can conclude that the test word was activated by the anaphor.

TABLE 1

AN EXAMPLE OF THE PARAGRAPHS USED IN THE EXPERIMENTS

Critical word (referent of anaphor): *burglar*
Companion word: *garage*
Control word: *bottles*
Sentence 1: A burglar surveyed the garage set back from the street.
Sentence 2: Several milk bottles were piled at the curb.
Sentence 3: The banker and her husband were on vacation.
Sentence 4: (Version 1, anaphor): The$_1$ criminal$_2$ slipped$_3$ away$_4$ from the$_5$ streetlamp.$_6$
Sentence 4 (Version 2, no anaphor): A$_1$ cat$_2$ slipped$_3$ away$_4$ from the$_5$ streetlamp.$_6$

Note. The number subscripts represent test positions.

EXPERIMENT 1

The first concern is to see if one can get any priming at all in this paradigm. Hence we first used only the referent (*burglar*) as a test word. Because it is already known that the referent is activated at the end of the clause (Chang, 1980; McKoon & Ratcliff, 1980), we tested three positions within the clause, positions 2, 3, and 4. These positions represent 250, 500, and 750 milliseconds after presentation of the anaphor or the control word. If the activation process is rapid and automatic, priming should be obtained at all of these positions including position 2 (250 milliseconds).

Method

Subjects. Twenty-four Dartmouth students participated for extra credit in an introductory psychology course.

Materials. The paragraphs from McKoon and Ratcliff's (1980) study were used. There were 60 paragraphs like that presented in Table 1. The first sentence of each paragraph contained a critical word as the first noun (e.g., *burglar*). The second and third sentences did not repeat or refer to this word. (However, the paragraphs were not unnatural.) There were two possible fourth sentences for each of these paragraphs. In version one, the category to which the critical word belonged was the first noun (e.g., *criminal*). In the second version, the first noun was a word not related to the critical word (e.g., *cat*). Except for the first noun and its associated determiner (*a, the, some*, etc.), the two versions of the fourth sentence were identical, and they were always single-clause sentences. In addition to these experimental paragraphs there were 75 filler paragraphs which varied in length from two to six sentences. Forty of these were to act as negatives (trials in which the test word was not in the paragraph) and 35 as positive fillers whose purpose was to prevent subjects from developing hypotheses regarding the length and structure of the experimental paragraphs and the position of the test word.

Procedure. Subjects were tested in groups of either one or two in a single 50-minute session. Presentation of all materials was controlled by a microcomputer driven by Dartmouth's time-sharing computer system.

Each subject received 10 practice paragraphs before receiving the 60 experimental and 75 filler paragraphs. Presentation of each paragraph began with an instruction for the subject to press the space bar on a CRT keyboard to initiate the paragraph. When the space bar was pressed, the first word of the paragraph appeared in the upper left part of the CRT screen. Then, 250 milliseconds later, the second word appeared to the right of the first in the same line, separated by a single space. The first word remained on the screen. The remaining words of the sentence then appeared, one at a time, every 250 milliseconds, in exactly the same way—separated by a space, with all previous words in the sentence remaining on the screen. If the words of a sentence could not be presented on a single line, the display carried over to the next line. After the final word of the sentence appeared, all the words remained on for an additional 600 milliseconds. Then the words disappeared, and 500 milliseconds later the next sentence was presented in the same way.

Subjects were told to read the paragraphs, and that, at some time during the presentation of one of the sentences of each paragraph, a single test word would appear. The test word was presented in capital letters (unlike the words of the paragraphs which were in normal case) and was underlined with a row of asterisks. It appeared immediately to the right of the last presented word on the same line, and when it appeared all the words of the current sentence disappeared.

Subjects were instructed to press the "/" key on the keyboard if the test word was in the paragraph, the "Z" key if it was not, and to respond quickly and accurately. If the subject made an incorrect response, the word ERROR appeared on the screen, and if not, the response time was presented. Following this feedback, the signal to begin the next paragraph came on. With minimal practice, subjects adapted easily to the reading and test procedures.

The conditions of interest were represented in the 60 experimental paragraphs. For these paragraphs, the fourth sentence was either version one, containing an anaphor (*criminal*) for the critical word (*burglar*), or version two, containing a control word. In addition, the test word, which was always the critical word (*burglar*), was presented in either test position 2 (immediately after the anaphor or its control), position 3 (one word later), or position 4 (two words later). Thus, the test word appeared either 250, 500, or 750 milliseconds after the anaphor or its control.

The order of presentation of the experimental and filler paragraphs was random, and this order was changed for every second subject.

The six experimental conditions (2 versions of the paragraphs × 3 test positions) were crossed with six groups of subjects (four per group) and six sets of paragraphs (ten per set) in a Latin square design. Thus, each subject saw ten paragraphs in each condition, and the condition associated with each paragraph set was rotated across subject group.

Results and Discussion

For analyses of the response times, only correct responses were included. Analysis of variance was performed on the means for each condition for each subject (F_1 statistic) and the means for each condition for each paragraph (F_2 statistic). (This was true of all subsequent analyses in all experiments.) The means of the subject by condition means are presented in Table 2 along with the error rates.

As can be seen, response times were more rapid when the fourth sentence mentioned an anaphor to the test word than when it did not, $F_1(1,23) = 4.5$, $p < .05$; $F_2(1,59) = 5.2$, $p < .05$. This effect occurred at all testword positions as indicated by the means and by the lack of an interaction between test word position and paragraph version, $F_1 < 1$; $F_2 < 1$. The only significant effect in the error rates was a difference over test word position in the subjects analysis, $F_1(2,46) = 4.0$, $p < .05$.

TABLE 2
RESULTS FROM EXPERIMENT 1

Paragraph version	Test position 2	3	4
	Test word—Referent (*burglar*)		
1 (anaphor)	585 (10.4)	588 (10.4)	591 (6.3)
2 (control)	602 (8.8)	609 (14.2)	626 (7.5)

Note. Response time is in milliseconds and error rates (parentheses) are percentages. The average error *MS* for the subject mean response times was 4381.

Clearly the anaphor activated the critical word, and did so rapidly, possibly within 250 milliseconds and definitely within 500 milliseconds. However, at this point it cannot be concluded that the critical word was activated because it is the referent of the anaphor. It could have been activated solely because it is semantically related to the anaphor. In the second experiment we included a companion (*garage*) as a test word in addition to the referent (*burglar*). If the companion is activated in the same way as the referent, it would be difficult to argue that the activation is due to semantic priming rather than to the retrieval of antecedent information. This result would also support the position that propositions rather than concepts are retrieved when an anaphor is processed.

The second experiment used different test positions than the first, including one immediately before the anaphor. In this position, there should be no priming of either referent or companion because the only difference between the two versions of the paragraphs is that in version one, the word *the* has occurred whereas in version two, either the word *a* or the word *some* has occurred.

EXPERIMENT 2

Method

Subjects. Forty-eight subjects from the same population participated.

Materials and procedure. This experiment used the same materials and procedure as the previous one with two exceptions. First, the additional factor of test word type was introduced. The test word was either the first noun of the first sentence (the referent, *burglar*, as before) or a noun from the same proposition as the first noun (the companion, *garage*). Secondly, the test positions were position 1 (immediately before the anaphor or its control), position 3 (500 milliseconds after the anaphor), or position 5 (1250 milliseconds after). Thus there were 12 conditions (2 paragraph versions × 3 test word positions × 2 test words). Conditions were assigned to paragraph and subject groups using a 12 × 12 Latin square, so that each subject did five paragraphs in each condition, and so that, across the experiment, each paragraph participated equally in the conditions.

Results

The means of the subject by condition means are shown in Table 3. The largest effect was that of the test word. When the test word was the referent, response times were substantially shorter than when it was the companion, $F_1(1,47) = 56.7, p < .00001$; $F_2(1,59) = 44.0$, $p < .0001$. As in the first experiment, response times were faster for the version mentioning the anaphor than for the version mentioning the control word $F_1(1,47) = 5.8, p < .05$; $F_2(1,59) = 5.3, p < .05$. But, as would be expected, this effect interacted with test word position, $F_1(2,94) = 4.0, p < .05$; $F_2(2,118) = 3.1, p < .05$. In test position 1, before the anaphor or its control, there was no effect of sentence version. This result is not entirely trivial because the sentences containing the anaphor have the determiner "the" and the control versions have either "a" or "some" and this difference has occurred by test position 1. By test position 3, the effect of paragraph version is strongly established for both test word types, 41 milliseconds for the anaphor's referent (*burglar*) and 45 milliseconds for the companion (*garage*). This difference is reduced by test position 5 to 26 milliseconds for the referent (though still significant, $p < .05$), and is completely wiped out at this position for the companion (1 millisecond). However, we are not in a position to conclude that the patterns differ for the two test words because the triple interaction of test word, paragraph version, and test position was not significant. However, there was this triple interaction on the analysis of error rates, $F_1(2,94) = 4.9, p < .01, p < .01; F_2(2,118) = 7.0, p < .01$. This interaction can be easily seen in the error rates presented in Table 3. In test position 1, there is little or no effect of paragraph version for either test word. In position two, there are more errors associated with the control version than with the version containing the anaphor for both test words. However, by position 5, the patterns differ for the two test words. When the test word is the referent of the anaphor, the error rate is substantially lower when the anaphor is present (3.8%) than when it is not (12.1%). But when the test word is only in the same proposition as the referent of the anaphor, there is no advantage associated with mentioning the anaphor. In fact, we obtained more errors (20%) in the condition where the anaphor is present than in the control (16%) for this test position and word combination. The only other significant effect for errors was a clear superiority for the referent over the companion, $F_1(1,47) = 32.6, p < .0001; F_2(1,59) = 13.9, p < .001$.

TABLE 3
RESULTS OF EXPERIMENT 2

Paragraph version	Test position 1	3	5
	Test word—Referent (*burglar*)		
1 (anaphor)	653 (10.0)	641 (7.5)	651 (3.8)
2 (control)	650 (8.8)	682 (11.3)	677 (12.1)
	Test word—Companion (*garage*)		
1 (anaphor)	721 (13.8)	700 (14.2)	718 (20.8)
2 (control)	720 (16.2)	745 (19.2)	719 (16.2)

Note. Response times are in milliseconds and error rates (in parentheses) are percentages. The average error *MS* for the subjects analysis of response times was 7237.

Both the response times and error rates support the contention that the companion as well as the referent is activated by the anaphor and suggest that these concepts become activated at about the same time. This supports the hypothesis that antecedent information is retrieved in the form of propositions. However, the pattern of priming at 1250 milliseconds after the anaphor (position 5) qualifies this conclusion. The referent is still activated but the companion is not. One might argue that by this point the two concepts are being treated differently by the comprehension process.

The following two experiments are designed to provide more information on this and other issues.

EXPERIMENTS 3 AND 4

Experiment 3 was primarily a control experiment. The reduced response time and error rate associated with the paragraph version which mentioned the anaphor could be due to the retrieval of antecedent information as we claim, or it could be due to a general benefit that comes from processing an anaphor. If this latter claim were true then *any* test word, regardless of whether it is associated with the referent, should be primed by the anaphor. Thus, this experiment included a *control* test word which was not in the same sentence as the referent (e.g., *bottles*). We expected that this word would not be primed at all. In addition, the companion was included as a test word to gain more data on the issue of when this word ceases to be activated. This experiment used test positions 3 (500 milliseconds), 5 (1250 milliseconds), and 6 (end of sentence).

In Experiment 4, we tested the activation of the referent and the companion at position 2 (250 milliseconds after the anaphor), position 5 (1250 milliseconds after), and position 6 (at the end of the sentence). Experiment 1 provided tentative evidence that the referent is activated at 250 milliseconds. Experiment 4 should substantiate this result

and also determine whether the companion becomes activated this early. The test at position 5 is an attempt to replicate the unexpected pattern found in Experiment 2, in which the referent remained activated and the companion did not. Additional information on this question should also come from the test at position 6, the end of the sentence.

Method

Subjects. Each experiment used 48 subjects from the same population as before.

Materials and procedure. These experiments were the same as Experiment 2, except that the test words and test positions were different. For Experiment 3 the test words were either the control word (*bottles*) or the companion (*garage*) and the test positions were 3, 5, and 6. For Experiment 4 the test words were either the referent or the companion and the test positions were 2, 5, and 6.

When the test word was given at position 6, it actually appeared on the beginning of the next line following the 500 millisecond pause that normally occurred between sentences. On average, there were 6.5 words between the anaphor or control word and the end of the sentence. Thus, a test word in position 6 appeared an average of 2725 milliseconds after the anaphor or control word.

Results and Discussion

First consider the control experiment, Experiment 3. The means and error rates are shown in Table 4. Turning first to the control test word (*bottles*), it can be easily seen that there is no priming in either response times or error rates at any test position. Clearly, the anaphor did not activate the control word. Thus it seems that any activation that does result from the anaphor is not due to a general enhancement of the availability of all words in the paragraph.

The pattern for the companion test word (*garage*) is the same as that of the previous experiment. There is priming (25 milliseconds) at test position 3, which was reduced by position 5 (to 11 milliseconds). Furthermore there is no evidence of priming at the end of the clause (position 6). However, the pattern is not as clean as it was in Experiment 2, in which there was strong priming for the companion at position 3 and none at 5. In fact, an overall analysis of variance on response times did not yield any significant effects of the paragraph version variable, nor of any of its interactions with test word or test position. There were, however, reliable main effects of test word (control vs companion) on response times and error rates, all F's > 25, and of test position, F's > 10 for response times and F's > 4 for the error rates.

TABLE 4
RESULTS OF EXPERIMENT 3

Paragraph version	Test position 3	5	6
	Test word—Companion (*garage*)		
1 (anaphor)	709 (15.0)	736 (16.0)	752 (21.2)
2 (control)	734 (13.2)	747 (19.4)	753 (18.7)
	Test word—Control (*bottles*)		
1 (anaphor)	672 (7.5)	678 (10.8)	721 (10.4)
2 (control)	672 (6.7)	691 (8.7)	724 (9.6)

Note. Response times are in milliseconds and error rates (in parentheses) are percentages. The average error *MS* for the subjects analysis of response times was 6965.

Although the pattern of priming was similar to that of Experiment 2 for the companion, we would like more evidence before we claim that this word is activated at position 3 and not at position 5. The final experiment provided such evidence.

The means and error rates are presented in Table 5. As expected there was significant priming (the main effect of paragraph version) in both response times and error rates, all F's > 5. The effects of interest, though, concern the patterning of response time priming across test words and positions. The only interaction that was significant was that between paragraph version and test word, and only for the subjects analysis, $F_1(1,47) = 3.9, p = .05$, showing that priming is more evident in the referent

TABLE 5

RESULTS OF EXPERIMENT 4

Paragraph version	Test position 2	5	6
Test word—Referent (*burglar*)			
1 (anaphor)	643 (6.2)	657 (10.0)	661 (8.3)
2 (control)	684 (11.6)	675 (13.4)	696 (9.2)
Test word—Companion (*garage*)			
1 (anaphor)	705 (15.0)	716 (19.0)	762 (17.0)
2 (control)	737 (18.0)	717 (21.2)	748 (16.6)

Note. Response times are in milliseconds and error rates (in parentheses) are percentages. The average error *MS* for the subjects analysis of response times was 7357.

than in the companion. However, the pattern of priming for the test words across test positions is strikingly consistent with that of the earlier experiments. Both the referent and the companion were strongly activated at position 2, 41 milliseconds of priming for the referent and 32 milliseconds for the companion. Thus it appears that both of these concepts become activated quite early. By 1250 milliseconds after the anaphor (position 5), this priming is gone for the companion (1 millisecond) but is still present for the referent (18 milliseconds). The finding of no priming for the companion at position 5 replicates the findings of Experiments 2 and 3, and the presence of modest priming for the referent at this position goes along with the results of Experiment 2. Thus, it is reasonably well established that, although both the referent and companion become active very early and to an equal extent, the companion's activation dies away before that of the referent. This can also be seen in the test at the end of the sentence (position 6). Here the referent is still active (35 milliseconds) and the companion is definitely not (14 milliseconds in the wrong direction).

At this point we would like to turn to the analysis of the four experiments as a group, and particularly the last three. Our claim that priming varies as a function of the test word–test position combination received only a little statistical support. The key triple interaction was significant only in Experiment 2, and only for the error rates. (Various double interactions involving priming were significant, but these are not critical to the claim of different time courses of activation for the different words.)

Viewed together, the experiments reveal an interpretable and (as we shall argue) statistically significant triple interaction. Figure 1 reveals this interaction very clearly for response time by summarizing the amount of priming obtained in all four experiments. Each point in the figure represents an average of the mean response times from the experiments that included that condition. So, for example, the point for the referent at position 2 (29 milliseconds) is an average of the amount of priming obtained in Experiment 1 (17 milliseconds) and in Experiment 4 (41 milliseconds).

Leaving aside position 1, the difference between paragraph versions (priming) is relatively constant across positions for the referent, decreasing across positions for the companion, and never present for the control. Computing facilitation through the error rates reveals the same pattern. The priming effects (percentage of errors for the control paragraph minus that for the anaphor paragraph) for the referent were +1.8, +3.8, +1.2, +5.9, and +0.9 for positions 2–6, respectively. For the companion

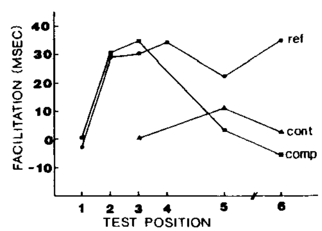

FIG. 1. The average amount of priming (RT for version 2 − RT for version 1) in milliseconds as a function of test word type and test position. Test position 1 is immediately before the anaphor or its control; position 2 is 250 milliseconds after the onset of the anaphor or its control; positions 3, 4, and 5 are 500, 750, and 1250 milliseconds after, respectively; and position 6 is after the sentence.

priming was only obtained at positions 2 and 3 (+3.0 and +1.6) and not at positions 5 and 6 (+0.3 and −1.4). Finally, there was no priming for the control word at any position (−0.8, −2.1, and −0.8 at positions 3, 5, and 6, respectively).

The question arises: Why was this clear pattern not always revealed in the statistics of the individual experiments? One reason is that the triple interaction, at least with response time, is small numerically. This is because it is not a cross-type of interaction. To illustrate the point, consider the eight response times corresponding to positions 3 and 5 for the referent and the companion in Experiment 2 (Table 3). For these data points the triple interaction is 14.5 milliseconds. Intuitively it seems larger. There is priming for the referent at both positions, but only for the companion at position 3. Nonetheless, the effect is small and so it is reasonable to conclude that the experiments individually lacked power to detect this sort of interaction.

So far it has been suggested that there is a triple interaction for response time when all experiments are considered together. Can this overall effect be tested for significance? Statistical procedures can be used; it is simply a question of their appropriateness. We will first describe our analysis and discuss the assumptions that must be made, and then present the results of it.

The basic idea is that, across the experiments, many of the conditions were tested two or three times. The variation in the effects for these conditions over the replications represents experimental error—the nonreplicability of the effects. If the effects are large relative to this error, then one can conclude that they are significant. Specifically, we performed an analysis of variance on the mean facilitation in milliseconds due to paragraph version (control version − anaphor version) over test positions and test words, in which the means were calculated for each experiment. Only the last three experiments were included. These numbers can be found in Table 6. As

TABLE 6

AMOUNT OF PRIMING OBTAINED IN EXPERIMENTS 2, 3, AND 4 AS A FUNCTION OF TEST WORD AND POSITION

Test words	Test positions 2 & 3	5	6
Referent	+41, +41	+26, +18	+35
Companion	+32, +45, +25	+1, +1, +11	+1, −14
Control	0	+13	+3

shown in the table the design had three levels of test word, three of test position, and a variable number of observations per cell. Data from position 1 (Experiment 2) were not included because it was never tested with the control word and it is largely uninteresting. Positions 2 and 3 were combined to increase the error degrees of freedom and because position 2 was never tested with the control word. Notice that analyzing the amount of priming rather than response time makes the critical triple interaction into a double one, that between test word and position.

Is it appropriate to use analysis of variance on these numbers? The main problem is that some numbers came from the same experiment and thus might be correlated, thereby violating the standard assumption of independent sampling of the error component of each score. However, because each experiment was identical in design, procedure, and materials, and used a large number of subjects (48) drawn from the same population there should be little dependence among the scores arising from particular experiments. Thus, by analyzing the data in this way we are making a strong assumption, but one which is likely to be true.

Turning to the results of the analysis, it was found that the test word and test position factors interacted significantly, $F(4,7) = 4.9$, $p < .05$, showing that the pattern of priming across positions differed for the test words. This difference did not come only from the inclusion of the control test word in the analysis. With the control word left out, the resulting 2 × 3 analysis still had a significant interaction, $F(2,7) = 5.0, p < .05$. Given this interaction we then tested individual condition means. First, it was ascertained by t tests (with the standard alpha level of .05 divided into nine equal parts) that priming is significantly different from zero at all test positions for the referent, only at positions 2 and 3 for the companion, and never for the control word. Second, the amount of priming for the referent and companion was compared at each position. There was no difference at positions 2 and 3, $t(7) = 1.0$, and significant differences at position 5, $t(7) = 2.5, p < .05$, and at position 6, $t(7) = 4.4, p < .01$.

In the final section we will evaluate the impact of our findings on theories about the retrieval of antecedent information in text processing.

General Discussion

The experiments presented here make three empirical points. First, an anaphor gives rise to very rapid activation of its referent, and second, it gives rise to equally rapid activation of other concepts in the same proposition as the referent. Third, the referent remains active through to the end of the sentence, but the other concepts do not.

These results allow theoretical insights into the processes involved in reading. The first point to be discussed concerns the speed of retrieval of antecedent information. We found that priming had reached its asymptotic value when the anaphor was presented only 250 milliseconds in advance of a test word. Thus it appears that the retrieval process begins immediately when the anaphor is encountered. A reasonable analogy is that the processing system is a pipeline of information in which the anaphor needs only a 250-millisecond advantage in onset time over the test word to give maximum facilitation to the referent and companion words. This speed of retrieval suggests an automatic process, as automatic would be defined by Posner and Snyder (1975; see also Ratcliff & McKoon, 1981) (although the other criteria for an automatic process, slow onset of inhibition and effects of probability manipulations, have not been examined). Certainly, our results are inconsistent with theories stating that interword semantic relations are only computed at such places as deep-structure clause boundaries (e.g., Fodor, Bever, & Garrett, 1974).

What exactly happens during the retrieval process? First, the superordinate anaphor (*criminal*) must make contact with a previously mentioned subordinate (*burglar*). An experiment by Garrod and Sanford (1977) suggests that this contact cannot be explained simply as the result of semantic priming, that is, as the result of the semantic relationship that exists between *criminal* and *burglar* independently of any text. Garrod and Sanford found that reading time for a sentence containing a superordinate anaphor (e.g., *the vehicle*) was slower if the referent was an unlikely example of the category (e.g., *tank*) than if it was a likely example (e.g., *bus*). However, they ruled out simple semantic priming between super- and subordinate words as an explanation for the effect because changing the superordinate so that it was not an anaphor (*a vehicle*) reduced the effect substantially.

With our results, we can also argue against simple semantic priming as an explanation for the activation of referents. In our experiments there are two ways in which the semantic overlap between anaphor and referent can be used. First there is the relationship between the referent, which is in the first sentence, and the subsequently occurring anaphor. Second, there is the same semantic relationship between the anaphor and the test word (when the test word is the referent). If facilitation is due to the first of these relationships then it is due to the retrieval of antecedent information. If it is due to the second then it is due to priming from the anaphor to the test word. We claim that the

obtained facilitation over at least the early test positions (2 and 3) is due to the retrieval of antecedent information because the companion was activated to the same extent as the referent. The speed and strength of the companion's activation can only be due to retrieval from the first sentence. Thus it is reasonable to assert the same for the referent. Specifically, our interpretation is that a definite anaphor activates quickly (possibly automatically) the propositions containing the referent or possible referents. The activation of these propositions is undoubtedly triggered in part by the semantic relationship between the anaphor and the potential referents. But it is clear that the facilitation, itself, is not due to a semantic relationship between the anaphor and the test word.

Next, we turn to the finding that the referent stays activated until the end of the sentence whereas the companion does not. So far the results have given evidence for the rapid activation of propositions containing potential referents. The next stage would logically be one in which the proper referent is selected. The data for the later test positions (5 and 6) support this view. The activation of nonreferents such as the companion has decayed, but the referent remains active, either because it is still being processed or because it has replaced the anaphor in the representation of the text. In either case it would be facilitated relative to a condition in which there was no anaphor at all.

Another possibility for the continued activation of the referent must be considered, however. Recall that we argued that facilitation at the early test positions could not be due to priming from the anaphor to the test word because the companion was activated to an equal extent as the referent. At the later positions, however, the companion is no longer active. Thus the possibility arises that the continued activation of the referent is due either to the semantic relationship between the anaphor and the test word or to the semantic relationship between the anaphor and its previously-mentioned referent.

A final point regarding the pattern of priming at the later positions is that our results are in partial disagreement with those of McKoon and Ratcliff (1980). They found that the companion was active when tested at the end of the sentence, about 2.5 seconds after the anaphor, contradicting our finding of no effect for the companion at position 5 (experiments 2, 3, and 4) and position 6 (experiments 3 and 4). However, McKoon and Ratcliff presented their paragraphs one sentence at a time and allowed subjects to control the pace. Thus subjects could have read the critical sentences, or at least the anaphors, more than once, and so kept the companion words active. This rereading is not possible with the word-by-word technique.

To conclude, we will briefly discuss the research method used in the experiments. Unlike previous recognition studies where words were tested at the ends of sentences (McKoon & Ratcliff, 1980; Chang, 1980), the present experiments tested words during the sentences. This was made possible by the presentation mode, one word at a time every quarter of a second, with words appearing next to one another across the screen. The benefit of combining unexpected recognition tests with word-by-word presentation is that it is possible to look at the state of activation of concepts in the reader's memory during the process of reading. Other on-line measures of comprehension such as reading time or eye fixation time are useful in that they describe the reader's behavior, but they cannot reveal directly what concepts are active in memory. On the other hand, later recall and recognition tests can well describe the reader's memory, but only after the fact of comprehension. Thus the activation technique provides a useful addition to the set of techniques for investigating the processes involved in reading.

The present method bears a resemblance to the lexical decision technique developed by Swinney (1979). In this technique subjects listen to spoken text and at unexpected times are given a letter string for a lexical decison. If reaction time to a certain word string is speeded, it can be concluded that a concept related to that word was activated by the comprehension of the spoken text. Thus the item recognition and lexical decision techniques may provide converging evidence on the activation functions of concepts during text comprehension. One technique, recognition, requires reference back to the text just processed, whereas the other, lexical decision, does not. Whether this difference leads to processing differences is a question for further research.

References

Anderson, J. R. *Language, memory, and thought.* Hillsdale, N.J.: Erlbaum, 1976.

Anderson, J. R., & Bower, G. *Human associative memory.* Washington, D.C.: Winston, 1973.

Bever, T. G., & Hurtig, R. R. Detection of a nonlinguistic stimulus is poorest at the end of a clause. *Journal of Psycholinguistic Research,* 1975, **4**, 1–7.

Chang, F. Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory and Cognition,* 1980, **8**, 58–64.

Foder, J. A., Bever, T. G., & Garrett, M. F. *The psychology of language.* New York: McGraw-Hill, 1974.

Garrod, S., & Sanford, A. Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior,* 1977, **16**, 77–90.

Kintsch, W. *The representation of meaning in memory.* Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In G. L. Nilsson (Ed.), *Memory: Processes and problems.* Hillsdale, N.J.: Erlbaum, 1978.

McKoon, G., & Ratcliff, R. The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior,* 1980, **19**, 668–682.

Posner, M., & Snyder, C. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition.* Hillsdale, N.J.: Erlbaum, 1975.

Ratcliff, R., & McKoon, G. Automatic and strategic priming in recognition. *Journal of Verbal Learning and Verbal Behavior,* 1980, **20**, 204–215.

Rumelhart, D., Lindsay, P., & Norman, D. A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

Swinney, D. Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior,* 1979, **18**, 645–660.

(Received March 1, 1982)