The Role of Recognition Memory in Anaphor Identification

Stephen Dopkins

Catherine Trinh Ngo

The George Washington University

RUNNING TITLE: Memory and Anaphor Identification

Please address correspondence to:
Stephen Dopkins
Psychology Department
George Washington University
2125 G Street, NW
Washington, DC 20052
202-994-6315 (voice)
202-994-1602 (fax)
dopkins@gwu.edu

The Role of Recognition Memory in Anaphor Identification

Abstract

In studies of anaphor comprehension, the capacity for recognizing a noun in a sentence decreases following the resolution of a repeated-noun anaphor (Gernsbacher, 1989). In studies of recognition memory, the capacity for recognizing a noun in a scrambled sentence decreases following the recognition that another noun has occurred before in the scrambled sentence (Dopkins & Ngo, 2002). The results of the present study suggest that these two phenomena reflect the same recognition memory process. The results suggest further that this is not because participants in studies of anaphor comprehension ignore the discourse properties of the stimulus materials and treat them as lists of words upon which memory tests are to be given. These results suggest that recognition processes play a role in anaphor comprehension and that such processes are in part the means by which repeated-noun anaphors are identified as such.

Extensive memory resources are brought to bear in the comprehension of a discourse. Shorter-term memory supports the on-line computation of discourse meaning (Just & Carpenter, 1992; Almor, 1999). Longer-term memory provides information necessary for the drawing of inferences and preserves the content of the discourse once it is comprehended (Graesser, Singer & Trabasso, 1994; Kintsch, 1988). Much remains to be learned about the memory processes that are involved during comprehension. We still do not understand how these processes operate and how they are related to the memory processes that are studied with traditional memory tasks (Glenberg & Kruley, 1992; Gordon, Hendrick, & Foster, 2000; Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002).

The present study focused on a memory process that is associated with *anaphor* comprehension. Each time an entity is mentioned following its introduction in a passage, the entity is identified by means of an anaphoric expression. When such an expression is encountered, an *antecedent* must be located in the preceding material. When this is accomplished, the anaphor is said to be *resolved* (Cacciari, Carreiras, & Cionini, 1997; Garnham, 1985, 1987, 2001; Garnham, Oakhill, & Cain, 1997; Garrod, Freudenthal, & Boyle, 1994; MacDonald & MacWhinney, 1995).

Memory processes are implicated in anaphor resolution, by virtue of the fact that the capacity for recognizing words from a passage changes with the resolution of an anaphor in the passage. The anaphor's antecedent is recognized more quickly and/or more accurately at this point (Cloitre & Bever, 1989; Corbett & Chang, 1983; Dell, McKoon, & Ratcliff, 1983; Gernsbacher, 1989; O'Brien, 1987; O'Brien, Duffy, & Myers, 1986; O'Brien, Plewes, & Albrecht, 1990). At the same time, other words in the passage are recognized less quickly and/or less accurately (Gernsbacher, 1989; McDonald & MacWhinney, 1990; Nordlie, Dopkins, & Johnson, 2001). The present study explored the process underlying the latter phenomenon.

The phenomenon came to light in a study by Gernsbacher (1989). The stimulus materials for the

critical trials in the study were short, two-clause, sentences. In the first clause of each sentence, two characters were introduced, identified with proper nouns. At the beginning of the second clause, a repeated-noun anaphor occurred whose antecedent was the proper noun that had been used to identify one of the characters (e.g. Ann predicted that Pam would lose the track race, but Pam came in first very easily). Participants performed a *recognition probe task* using these sentences as stimulus material. On each trial, the participant read a sentence, word by word, and responded to a recognition probe that was presented either before or after the anaphor. A striking pattern of results was observed when the test word was the proper noun that was not the anaphor's antecedent. The *non-antecedent* was recognized more slowly after than before the anaphor.

Gernsbacher attributed her results to a process associated with anaphor comprehension. She proposed that the function of the process is to suppress *competitors* to the antecedent of a repeated-noun anaphor and thereby promote retrieval of the antecedent. In other words, the function of the process is to suppress words that might potentially be the antecedent but actually are not. As a consequence of the process, competitors are subject to a recognition decrement following the resolution of a repeated-noun anaphor (Gernsbacher, 1989, 1990) Gernsbacher's interpretation of her results may warrant closer scrutiny in the light of some recently-reported results.

Dopkins and Ngo (2002) observed something similar to Gernsbacher's recognition decrement using a procedure that followed Gernsbacher except that the order of the words in the stimulus sentences was scrambled. On each trial, the participant read a scrambled sentence, responded to a recognition probe, and indicated whether the sentence contained any repeated words. In the crucial condition, the test word was a noun from the sentence and the last word in the sentence was another noun that was repeated from earlier in the sentence. The test word was recognized less well in this condition than when 1) an adverb was inserted in place of the repeated noun, 2) a new noun was inserted in place of the repeated noun, or 3) the sentence ended immediately before the repeated noun.

Dopkins and Ngo (2002) concluded 1) that participants, perhaps under pressure from the requirement that word-repetition be monitored, recognized the repeated noun as having occurred before in the scrambled sentence, and 2) that the test noun was recognized less well as a consequence of this recognition judgement. Dopkins and Ngo inferred that the process underlying their results could not be specific to anaphor comprehension because their stimulus items had no discourse properties. Instead the process must be general to recognition memory, and must operate as follows: When a word (here, the repeated noun) is recognized as having occurred in a memory set (here, the scrambled sentence), other words in the memory set (here the test noun) are subject to a recognition decrement.

Dopkins and Ngo (2004) demonstrated that the Dopkins and Ngo (2002) recognition decrement generalizes to a traditional memory paradigm. On each trial, the participant read a scrambled sentence and made recognition judgments to two common nouns in succession. The second test word was recognized less well when the first test word came from the sentence than when it did not.

The Dopkins and Ngo (2002) recognition decrement may have interesting implications for our understanding of the Gernsbacher recognition decrement and of anaphor comprehension in general. According to this line of thinking, the Gernsbacher and Dopkins and Ngo phenomena reflect the same underlying process. Thus, Gernsbacher's phenomenon reflects a memory rather than a comprehension process. This is because the comprehension of a repeated-noun anaphor depends on the general capacity for recognition memory, as follows: In order for any anaphor to be resolved, it must first be identified as an anaphor. An initial clue that a noun should be identified as an anaphor emerges by virtue of the fact that the noun is recognized as having appeared previously in the current passage. When this happens, further processing determines whether the noun is in fact an anaphor. As a consequence of the act of recognition, other words in the passage are subject to a recognition decrement, through a process such as Dopkins and Ngo proposed.

In order to advance the forgoing interpretation, we must deal with two objections. According to

the first objection, the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena do not reflect the same underlying process. On one hand, the Gernsbacher recognition decrement reflects an anaphor comprehension process. On the other hand, the Dopkins and Ngo recognition decrement reflects a memory process. Under this view, then, the Dopkins and Ngo phenomenon has no relevance for the Gernsbacher (1989) phenomenon and thus no relevance for anaphor comprehension.

According to the second objection, the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect the same underlying process, but not for the reasons suggested earlier. Under this account, participants in recognition probe tasks ignore the discourse properties of the stimulus materials and treat them as simple lists of words upon which memory tests are to be given. Thus, participants in the Gernsbacher and Dopkins and Ngo tasks process the stimulus materials for those tasks in much the same way. As a consequence, the Gernsbacher and Dopkins and Ngo recognition decrements reflect the same recognition process. In neither case, however, do the phenomena have much to teach us about anaphor comprehension. Under this view, then, the Dopkins and Ngo (2002) phenomenon has relevance for the Gernsbacher (1989) phenomenon but no relevance for anaphor comprehension.

Notice that this second objection, if valid, has ramifications beyond interpretation of the recognition decrements of Gernsbacher (1989) and Dopkins and Ngo (2002). At issue is the value of recognition probe results as a means of studying comprehension processes. If this objection is valid, interpretations of other recognition probe results are called into question. In fact, interpretations of recognition probe results have been criticized along these lines previously (Gordon, et al., 2000). Thus, providing evidence counter to this second objection would reinforce the interpretation of previous results obtained with recognition probe tasks.

The present study sought to advance the view that the Gernsbacher recognition decrement reflects a memory rather than a comprehension process and that this is because the processing of a repeated-noun anaphor involves the general capacity for recognition memory. Toward that end, the

study sought evidence against the two objections raised earlier with respect to that view. Experiments 1A and 1B sought evidence against the view that the Gernsbacher and Dopkins and Ngo phenomena reflect distinct processes. Experiments 2 and 3 sought evidence against the view that participants in recognition probe tasks ignore the discourse properties of the stimulus materials.

Experiments 1A and 1B

These experiments sought evidence against the view that the Gernsbacher and Dopkins and Ngo phenomena reflect distinct processes. The strategy was to explore the *extent* of the Gernsbacher phenomenon. Gernsbacher has shown that other words in a passage are subject to a recognition decrement following the processing of a repeated-noun anaphor. What about words that are not part of the passage but are part of the experience of comprehending the passage? Will such words also be subject to a recognition decrement? Our expectations differ depending on whether or not the Gernsbacher and Dopkins and Ngo phenomena reflect distinct processes.

More concretely, consider the following modification of the recognition probe task: On each trial, the participant reads a short passage, responds to two recognition probes, and answers a comprehension question. The first recognition probe occurs during the presentation of the body of the passage. After the participant responds to this probe, the passage resumes. At a subsequent point, the passage is truncated and the second probe occurs (Table 1 shows a sample passage with recognition probes inserted). In the After condition, the last word of the passage is a repeated-noun anaphor. In the Before condition, the last word in the passage is the word that preceded the anaphor in the After condition. In other words, the passage is truncated one word earlier here. The same test word, a noun that does not appear in the passage, is presented at the first and second probe point. At each probe point, the participant indicates whether or not the test word has appeared *during the current trial*. When the test word appears at the first probe point, the correct response is "no," because the word has not previously appeared during the current trial. When the test word appears at the second probe point,

the correct response is "yes," because the word has previously appeared during the current trial – in the first recognition probe. The data of interest are the response times and error rates for the second recognition probe. Assume that the participant treats the test word as part of the experience of reading the passage but not part of the passage. Will the test word be subject to a recognition decrement following the processing of the repeated-noun anaphor that ends the passage in the After condition? In other words, will the test word be recognized less well in the After than the Before condition?

Insert Table 1 about here

If the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect distinct processes, then the test word should not be subject to a recognition decrement. According to this view, the Gernsbacher phenomenon reflects a comprehension process, the function of which is to suppress competitors of a repeated-noun anaphor's antecedent. As a consequence of the process, competitors are subject to a recognition decrement following the resolution of such an anaphor (Gernsbacher, 1989). Because the test word is not part of the passage, it is not a competitor of the anaphor's antecedent. Thus, this word will not be subject to a recognition decrement following the processing of the anaphor.

By contrast, if the Gernsbacher and Dopkins and Ngo phenomena reflect the same memory process, the test word may be subject to a recognition decrement following the processing of the anaphor. To see this, we must consider two different scenarios under which the proposed memory process may operate.

Under the first scenario, the participant ignores the discourse properties of the passage and treats it as a simple list of words upon which a memory test is to be given. The participant recognizes the anaphoric noun as having appeared before in the passage/list. When this happens, a recognition decrement occurs. Under the second scenario, the participant attends to the discourse properties of the passage. The participant is consequently alert for the presence of anaphors. Toward this end, the

participant is sensitive, at some level, to the occurrence of nouns that have appeared before in the passage. When the participant recognizes the anaphoric noun as having appeared before, a recognition decrement occurs. (Clearly, the second scenario is preferred in the current project. Later experiments will attempt to establish that this is the correct scenario).

Under either scenario, a recognition decrement occurs when the anaphoric noun is recognized as having appeared before in the passage. At this point, other words in the memory set associated with the recognition operation are recognized less well than would otherwise be the case. The memory set consists roughly of the words from the current passage. Crucially, however, the memory set may not be strictly limited to words from the current passage. Given that the memory set is associated with a memory rather than a comprehension process, the memory set may include words that are part of the experience of comprehending the passage but not part of the passage per se. Thus, the test word may be included in the memory set. In this case, the word will be subject to a recognition decrement following the processing of the anaphoric noun.

Experiments 1A and 1B used a recognition probe task with crucial trials such as were described above to seek evidence against the view that the Gernsbacher and Dopkins and Ngo phenomena reflect distinct processes. The object was to find out whether the test word was recognized less well in the After than the Before condition. If the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect distinct processes, then such a result should not occur. On the other hand, if the Gernsbacher and Dopkins and Ngo phenomena reflect the same memory process, then such a result may occur.

Experiments 1A and 1B differed as follows. In Experiment 1A, participants were tested in a single session with the procedure described above. Experiment 1B sought to reinforce the results of Experiment 1A, while at the same time taking measures to insure that relative novelty of the procedure for that experiment did not compromise the results. Specifically, participants were tested in an initial training phase, in which they became familiar with the procedure, and returned subsequently for the main

test phase.

Experiment 1A

Method

Participants. The participants were 20 students at the George Washington University. They received extra credit in a psychology course in exchange for their efforts. Some of the participants produced rather high error rates to the second recognition probe, probably because they did not understand the task instructions. Participants were replaced if their overall error rate for responses to the second probe was greater than .5. Four participants were replaced under this criterion.

Design. The position of the second recognition probe was manipulated within participants and items.

Materials. Thirty-six experimental passages were used. Each consisted of two sentences and mentioned three characters, with proper nouns being used to identify all of the characters. A sample passage is presented in Table 1. Two of the characters were introduced in the first sentence and the third was introduced in the first clause of the second sentence. The first recognition probe was presented at some point during the first sentence or the first clause of the second sentence. In the After condition, the name of the first character reappeared anaphorically in the second sentence. The second recognition probe occurred immediately after the anaphor, and the passage was discontinued. In the Before condition, the second probe occurred immediately after the word that preceded the anaphor in the After condition, and the passage was discontinued. The test word for both the first and second probes was a common noun that did not appear in the passage and that had no semantic relationship to the content of the passage. The experimental passages were rotated through the After and Before conditions in such a way as to create two presentation sets.

The experimental passages were randomly intermixed with 63 filler passages. With these passages, a different test word was used for the first and second recognition probe. For twenty of the

filler passages, the correct response to both test words was “yes.” Of these passages, the test words were both nouns for fourteen passages, a noun and a verb for three passages, and a verb and a noun for three passages. For twenty-three of the filler passages, the correct responses to the test words were “yes” and “no” respectively. Of these passages, the test words were both nouns for seven passages, a noun and a verb for six passages, a verb and a noun for eight passages, and both verbs for two passages. For nine of the filler passages, the correct responses to the two test words were “no” and “yes” respectively. Of these passages, the test words were both nouns for six passages, a noun and a verb for one passage, and a verb and a noun for two passages. For eleven of the filler passages, the correct response to both test words was “no.” Of these passages, the test words were both nouns for eight passages, a noun and a verb for two passages, and a verb and a noun for one passage. Each sentence of a given experimental and filler passage began with a capital letter and ended with a period.

Procedure. Each participant was randomly assigned to a presentation set. The passages were presented on a computer monitor, according to a procedure derived from Gernsbacher (1989). The participant started the presentation of each passage by pressing the space bar of the computer. The passage was presented word by word, with each successive word appearing alone in the middle of the screen for an amount of time determined by the following formula: Presentation Time = (300 ms) + ((16.667 ms) x (number of letters)). An inter-word interval of 150 ms occurred between each successive pair of words. Gernsbacher used the same procedure except that the constant base was 450 rather than 300 ms and there was no inter-word interval. Thus, the difference between the two procedures is that the participant saw each word for an extra 150 ms in the Gernsbacher procedure whereas he/she saw an extra 150 ms of blank screen between words in present procedure. Notice, however, that the rate of presentation was the same in the two cases.

At the first probe point, the test word appeared at the top of the screen, in capital letters. After the participant responded, the test word disappeared and the passage continued until the second probe

occurred. After the participant responded to the second probe, the test word disappeared and two yes/no comprehension questions appeared, in succession, at the bottom of the screen. Each question remained on the screen until the participant responded. The participant was instructed to respond as quickly as possible to the recognition probes, without sacrificing accuracy, and to respond as accurately as possible to the comprehension questions. For each recognition probe, the participant was instructed to make a positive response if the test word had appeared *during the current trial*. The participant made positive and negative responses by pushing the “B” and “N” keys, respectively.

Notice that the presentation procedure differed from Gernsbacher's in that the passage was discontinued after the second probe was presented. It has been shown in other research that this discontinuation procedure produces similar results to the standard procedure (Nordlie, et al., 2001).

Results and discussion

Performance on the first recognition probe was good and did not differ in the Before and After conditions, Response time. Before: 1238, SEM = 79; After: 1197, SEM = 73; $F1(1,18) = 1.08$, $MSe = 15{,}334$, $F2(1,34) < 1$; Error rate. Before: .07, SEM = .020; After: .07, SEM = .015; $F1(1,18) < 1$, $F2(1,34) < 1$. (F1 and F2 report the results of tests against the participant and item variability, respectively. All tests were conducted at the .05 level).

Analyses of the data for the second recognition probe were restricted to the data for trials upon which the response to the first probe was correct (i.e. "no"). Response time for the second probe did not differ in the Before and After conditions, Before: 1265, SEM = 70; After: 1314, SEM = 84; $F1(1,18) = 1.15$, $MSe = 20{,}897$, $F2(1,34) < 1$. Error rate for the second probe was greater in the After than the Before condition, Before: .25, SEM = .028; After: .33, SEM = .035; $F1(1,18) = 5.79$, $MSe = .01$, $F2(1,34) = 5.44$, $MSe = .02$.

A test word that appeared while a passage was being processed, but that was not part of the passage, was recognized less well after as opposed to before the processing of a repeated-noun

anaphor. The key results showed up in the error-rate rather than the response-time data. Crucially, a speed-accuracy tradeoff did not occur.

These results argue against the view that the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect different processes. According to this view, the Gernsbacher recognition decrement reflects the operation of a comprehension process, the function of which is to suppress competitors of a repeated-noun anaphor‘s antecedent. A word that is not part of the current passage is not a competitor of an anaphor‘s antecedent. Such a word should therefore not be suppressed with the processing of the anaphor. Thus, under the assumption that participants treated the test word as distinct from the passage, the test word should not have become more difficult to recognize with the processing of the anaphor.

These results argue for the view that the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect the same memory process. This view would accommodate the results as follows: When the anaphoric noun occurred in the present experiment, it was recognized as having appeared before in the current trial. Other words in the memory set associated with the recognition operation became more difficult to recognize as a consequence. The memory set consisted roughly of the words from the passage in which the noun appeared. The memory set also included words that were part of the experience of comprehending the passage, including the test word. Thus, the test word became more difficult to recognize with the processing of the anaphoric noun.

Experiment 1B

We must be somewhat wary of the results of Experiment 1A because the experimental task was somewhat novel. Given the novelty of the task, several concerns must be addressed. One concern is that the non-standard procedure induced participants to process the passages as simple lists of words rather than as passages of discourse. This concern will be addressed in Experiment 2. Another concern is that participants did not truly understand the ground-rules of the task and that a substantial amount of noise was consequently injected into the data. Experiment 1B sought to allay the latter concern. Each

participant was tested in a two-phase session. During the training phase of the session, the participant practiced performing the task of Experiment 1A until he/she could perform it to a high standard of accuracy. During the test phase, the participant was tested as in Experiment 1A.

Method

Participants. The participants were 30 students at the George Washington University. They received extra credit in a psychology course in exchange for their efforts.

Materials. Thirty-three passages were used in the training phase. Sixty-six passages were used in the test phase. Of the sixty-six test-phase passages, thirty-six were experimental passages and the rest were fillers. The experimental passages were the same as were used in Experiment 1A. The passages for the training phase and the filler passages for the test phase were the same as served as fillers in Experiment 1A.

For eight of the filler passages, the correct response to both test words was "yes." Of these passages, the test words were both nouns for two passages, a noun and a verb for two passages, a verb and a noun for two passages, and both verbs for two passages. For eight of the filler passages, the correct responses to the two test words were "yes" and "no" respectively. Of these passages, the test words were both nouns for two passages, a noun and a verb for two passages, a verb and a noun for two passages, and both verbs for two passages. For six of the filler passages, the correct responses to the two test words were "no" and "yes" respectively. Of these passages, the test words were a noun and a verb for three passages, and a verb and a noun for three passages. For eight of the filler passages, the correct response to both test words was "no." Of these passages, the test words were both nouns for two passages, a noun and a verb for two passages, a verb and a noun for two passages, and both verbs for two passages.

Procedure. The experimental session was two hours in duration. The training phase began at the start of the session and continued for as long as was necessary until the participant mastered the

task. Specifically, the participant was required to cycle repeatedly through the 33 training passages until he/she achieved a criterion level of 85% correct across the 33 recognition probes. 26 participants achieved the criterion on their first cycle through the passages. Four participants had to repeat the cycle once. Participants were dismissed upon completing the training phase and returned later for the test phase, which occurred during the final 30 minutes of the session. The trial-to-trial procedure during the training and test phase was as in Experiment 1A.

Results and discussion

Performance on the first recognition probe was good. Response time did not differ in the Before and After conditions, Before: 997, SEM = 47; After: 1010, SEM = 42; $F1(1,28) < 1$, $F2(1,34) < 1$; The error rate for the first probe was greater in the After than the Before condition when tested against the variability due to participants but not when tested against the variability due to items, Before: .03, SEM = .012; After: .04, SEM = .015; $F1(1,28) = 6.89$, $MSe = .001$, $F2(1,34) = 2.82$, $MSe = .001$. Although responses to the first probe tended to be more accurate in the Before condition, these differences do not compromise the results for the second probe because analyses of data for the second probe were limited to trials on which the response to the first probe was correct.

With respect to the second probe, response time was greater in the After than the Before condition, Before: 1041, SEM = 22; After: 1102, SEM = 26; $F1(1,28) = 2.90$, $MSe = 14{,}241$ (one-tailed), $F2(1,34) = 3.87$, $MSe = 16{,}902$. The error rate for the second probe was greater in the After than the Before condition when tested against the variability due to items but not when tested against the variability due to participants, Before: .11, SEM = .016; After: .14, SEM = .015; $F1(1,28) = 1.73$, $MSe = .006$, $F2(1,34) = 5.15$, $MSe = .006$.

As in Experiment 1A, a test word that appeared while a passage was being processed but that was not part of the passage was recognized less well after as opposed to before the processing of a repeated-noun anaphor. Whereas the effect showed up in the error-rate data in Experiment 1A, it

showed up in the response-time as well as the error-rate data here. These results are again inconsistent with the view that the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect different processes. Performance was more accurate in the present experiment than in Experiment 1A. By implication, the training procedures of Experiment 1B were effective in reducing the novelty of the task.

Experiments 2 and 3

The results of Experiments 1A and 1B suggest that the Gernsbacher (1989) and Dopkins and Ngo (2002) recognition decrements reflect the same process. These results may have application to our understanding of anaphor comprehension. One possible implication of the results is that the identification of a repeated-noun anaphor as such depends on recognition memory processes. According to this *anaphor relevant* interpretation, an initial clue to the effect that a noun is a repeated-noun anaphor emerges by virtue of the fact that the noun is recognized as having appeared before in the current passage. As a byproduct of the act of recognition, other words in the passage are subject to a recognition decrement, through a process such as Dopkins and Ngo proposed.

Before proceeding down this theoretical path, we must consider a less interesting interpretation of the results of Experiments 1A and 1B. According to this *anaphor irrelevant* interpretation, participants in a recognition probe task ignore the discourse properties of the stimulus materials for the task and treat them as simple lists of words upon which memory tests are to be given. In other words, participants in the Gernsbacher and Dopkins and Ngo (2002) tasks treat the stimulus materials in essentially the same way. As a consequence of this, the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect the same memory process. Participants in the Gernsbacher task are not externally motivated to recognize the anaphoric noun in the way that participants in the Dopkins and Ngo task are motivated to recognize the repeated noun. Nonetheless, participants in the Gernsbacher task do recognize the anaphoric noun, perhaps because recognition is an automatic process. The result is a recognition decrement similar to the one that occurs in the Dopkins and Ngo task. This recognition

decrement tells us nothing about anaphor comprehension, however. Experiments 2 and 3 sought evidence against the anaphor irrelevant interpretation.

Experiment 2

Experiment 2 sought to recreate the task of Experiments 1A and 1B and to show that, in performing this task, participants did not ignore the discourse properties of the stimulus materials. On each trial of the experiment, the participant read an initial piece of text, responded to an initial recognition probe, read a second piece of text, responded to a second recognition probe, and responded to a comprehension question. At each probe point, the participant indicated whether the test word had appeared previously *during the current trial*.

In the Replication and One Sentence conditions, the two pieces of text constituted two parts of a single sentence (A sample item is shown in Table1). In the Replication/After and One Sentence/After conditions, the last word of the sentence was a repeated-noun anaphor with an antecedent in the first part of the sentence. In the Replication/Before and One Sentence/Before conditions, the last word in the sentence was the word that preceded the anaphor in the Replication/After and One Sentence/After conditions. In the Replication condition, the same test word, a noun that did not appear in the sentence, was presented at the first and second probe point. In the One Sentence condition, the first test word was a noun that did not appear in the sentence and the second test word was a noun that appeared in the first part of the sentence. In the Two Sentence condition, each piece of text was a distinct sentence and no relationship existed between the sentences. In the Two Sentence/After condition, the last word in the second sentence was a repeated-noun anaphor with an antecedent in that sentence. In the Two Sentence/Before condition, the last word in the second sentence was the word that preceded the anaphor in the Two Sentence/After condition. In the Two Sentence condition, the first test word was a noun that did not appear in either sentence and the second test word was a noun that appeared in the first sentence.

The data of interest were the response times and error rates for the second test word. The object was to test for recognition decrements in the Replication, One Sentence, and Two Sentence conditions. More concretely, the object was to find out, in each of these conditions, whether the second test word was recognized more poorly in the After than the Before condition. A recognition decrement was expected in the Replication condition. This condition recreates the situation of Experiments 1A and 1B, in which recognition decrements were observed. A recognition decrement was also expected in the One Sentence condition. This condition recreates the situation of Gernsbacher's (1989) study, in which recognition decrements were also observed.

The question of interest was whether a recognition decrement would occur in the Two Sentence condition. In this condition, the participant reads two unrelated sentences on a given trial, with the second test word coming from the first sentence. The anaphor irrelevant interpretation predicted a recognition decrement here. The anaphor relevant interpretation did not. The bases for the predictions were as follows. According to the anaphor irrelevant interpretation, participants will ignore the discourse properties of the stimulus materials. Thus, the two pieces of text should be similarly undifferentiated in the One and Two Sentence conditions. If a recognition decrement occurs in the One Sentence condition, a similar decrement should occur in the Two Sentence condition.

According to the anaphor relevant interpretation, participants will attend to the discourse properties of the stimulus materials. They will therefore be alert for the presence of anaphors. Participants will identify nouns as anaphors, in part, on the basis of recognition memory capacities. Normally, the fact that a noun has appeared before *in the current passage* constitutes a clue to the effect that the noun is a repeated-noun anaphor. When participants recognize the noun as having appeared before in the passage, other words in the passage are subject to a recognition decrement, through a process such as Dopkins and Ngo (2002) proposed. Words in previous passages are not subject to such a recognition decrement. In the Two Sentence Condition of the present experiment, the

two pieces of text constitute two unrelated sentences, each of which functions as a separate passage. Thus, participants will be alert, at some level, while reading each sentence of the Two Sentence condition, for the occurrence of nouns that have appeared before in that sentence. When participants recognize the anaphoric noun in the second sentence as having appeared before in that sentence, words in the second sentence will be subject to a recognition decrement. Words in the first sentence will not be subject to a recognition decrement. The second test word comes from the first sentence. Thus, the second test word will not be subject to a recognition decrement.

Method

Participants. The participants were 40 students at the George Washington University. They received extra credit in a psychology course in exchange for their efforts.

Design. Probe Position (Before, After) was manipulated within participants and items. With respect to Materials Type, the contrast between the One and Two Sentence conditions was manipulated within participants and items. The contrast between the Replication condition and the other two conditions was manipulated within participants and between items.

Materials. Forty-two experimental items were used. Fourteen of the items were used for the Replication condition. (A sample item is shown in Table 1). Three characters were introduced in the first part of the sentence, identified with common nouns. One of the character nouns reappeared as a repeated-noun anaphor at the beginning of the second part of the sentence. In the After condition, a second of the character nouns, which initially appeared at the beginning of the first part, reappeared as a repeated-noun anaphor at the end of the second part (and thus the end of the sentence). In the Before condition, the last word in the second part was the word that immediately preceded the anaphor in the After condition. The same test word, a noun that did not appear in the sentence, was presented at the first and second probe point.

Twenty-eight of the items were used for the One Sentence and Two Sentence conditions. The

One Sentence version of each of these items was configured in the same way as the Replication items. The first test word was a noun that did not appear in the sentence and the second test word was a noun that appeared toward the end of the first part of the sentence.

In the Two Sentence version of each of these items, the first sentence was identical to the first part of the sentence in the One Sentence condition. The second sentence had a completely different content than the first sentence. It mentioned a single character, identified with a common noun. In the After condition, the noun that had been used to identify the character appeared again at the end of the second sentence as a repeated-noun anaphor. In the Before condition, the last word in the second sentence was the word that immediately preceded the anaphor in the After condition. The first test word was a noun that did not appear in either sentence and the second test word was a noun that appeared toward the end of the first sentence.

Because the same items were used in the One and Two Sentence conditions, control was exerted over 1) the distance between the appearance of the second test word in an item and its presentation as recognition probe and 2) other possibly influential factors. It should be noted, however, that the single sentence in the One Sentence condition typically contained two anaphoric relations whereas the second sentence in the Two Sentence condition typically contained a single anaphoric relation.

The Replication and One/Two Sentence items were randomly intermixed to create a single presentation set. Four versions of this set were created by rotating the Replication items through the Before and After conditions and simultaneously rotating the rest of the items through the One Sentence and Two Sentence and the Before and After conditions.

The experimental items were randomly intermixed with 42 filler items. For six of the filler items, the stimulus material was a single sentence and the correct response to both test words was “yes.” Of these items, the test words were both nouns for one item, a noun and a verb for one item, a verb and a

noun for two items, and both verbs for two items. For six of the filler items, the stimulus material was a single sentence and the correct responses to the test words were “yes” and “no” respectively. Of these items, the test words were both nouns for one item, a noun and a verb for one item, a verb and a noun for two items, and both verbs for two items. For twenty of the filler items, the stimulus material consisted of a single sentence and the correct response to both test words was “no.” Of these items, the test words were both nouns for five items, a noun and a verb for five items, a verb and a noun for five items, and both verbs for five items. For ten of the filler items, the stimulus material consisted of two sentences and the correct response to both test words was “no.” Of these items, the test words were both nouns for two items, a noun and a verb for two items, a verb and a noun for three items, and both verbs for three items. For both the experimental and the filler items, each sentence began with a capital letter and ended with a period.

Procedure. Each participant was randomly assigned to a presentation set. The experiment was presented on a computer monitor, according to the same procedure as was used in Experiments 1A and 1B. As in the previous experiments, the participant was instructed to make a positive response to each test word if the test word had appeared *during the current trial*.

Results and discussion

Replication condition. The data for four participants were incomplete and were not included in any of the analyses. Performance on the first recognition probe was good and did not differ as a function of Probe Position (See Table 2). Response time: $F1(1,34) = 1.36$, $MSe = 47{,}715$, $F2(1,12) = 1.45$, $MSe = 12{,}928$. Error rate: $F1(1,34) = 1.59$, $MSe = .006$, $F2(1,12) = 2.04$, $MSe = .002$. As in Experiments 1A and 1B, analyses of the data for the second recognition probe were restricted to the data for trials upon which the response to the first probe was correct. Response time on the second probe did not differ in the Before and After conditions, $F1(1,34) < 1$, $F2(1,12) < 1$. The error rate was greater in the After than the Before condition, $F1(1,34) = 9.25$, $MSe = .038$, $F2(1,12) = 12.99$, $MSe =$

.011. The recognition decrement of Experiments 1A and 1B was replicated here. By implication, the Gernsbacher (1989) and Dopkins and Ngo (2002) recognition decrements reflect the same process.

One/Two Sentence conditions. Performance on the first recognition probe was good and did not differ as a function of the experimental conditions (See Table 2). Response time: Materials Type, $F1(1,36) < 1$, $F2(1,24) = 1.35$, $MSe = 36{,}959$; Probe position, $F1(1,36) < 1$, $F2(1,24) < 1$; Materials Type X Probe Position, $F1(1,36) = 1.52$, $MSe = 58{,}970$, $F2(1,24) = 1.93$, $MSe = 28{,}210$. Error rate: Materials Type, $F1(1,36) = 1.24$, $MSe = .01$, $F2(1,24) = 1.18$, $MSe = .008$ Probe position, $F1(1,36) < 1$, $F2(1,24) < 1$; Materials Type X Probe Position, $F1(1,36) < 1$, $F2(1,24) < 1$.

Performance on the second recognition probe was of greater interest (See Table 2). Response time did not differ in the One and Two Sentence conditions, $F1(1,36) < 1$, $F2(1,24) < 1$. Response time also did not differ in the Before and After conditions, $F1(1,36) = 2.48$, $MSe = 80{,}807$, $F2(1,24) = 2.15$, $MSe = 51{,}810$. The effects of Materials Type and Probe Position did not interact in the response-time data, $F1(1,36) = 1.58$, $MSe = 73{,}650$, $F2(1,24) < 1$.

The error rate was greater in the Two than the One Sentence condition, $F1(1,36) = 28.14$, $MSe = .019$, $F2(1,24) = 25.18$, $MSe = .015$. In addition, the error rate was greater in the After than the Before condition, $F1(1,36) = 20.72$, $MSe = .019$, $F2(1,24) = 16.98$, $MSe = .016$. Finally, the effects of Materials Type and Probe Position interacted in the error-rate data, $F1(1,36) = 3.35$, $MSe = .026$ (one-tailed), $F2(1,24) = 5.60$, $MSe = .011$. The difference between the error rate in the Before and the After conditions was significant in the One Sentence condition, $F1(1,36) = 18.25$, $MSe = .023$, $F2(1,24) = 20.92$, $MSe = .014$, but not in the Two Sentence condition, $F1(1,36) = 2.66$, $MSe = .022$, $F2(1,24) = 3.11$, $MSe = .013$.

A recognition decrement occurred for the second test word in the One Sentence condition. As in Experiment 1A, the recognition decrement showed up in the error-rate but not the response time data. Again, a speed-accuracy trade-off did not occur.

A recognition decrement did not occur in the Two Sentence condition. Although error rate was nominally greater in the After than the Before condition, the difference was not significant. Furthermore, response time was greater in the Before than the After condition. A post hoc analysis on the Before – After difference within the Two Sentence condition showed this difference to be significant when tested against the variability due to participants, $F(1,36) = 5.65$, $MSe = 55{,}059$.

These results are consistent with the anaphor relevant interpretation of the results of Experiments 1A and 1B. They are what would have been expected if participants had attended to the discourse properties of the stimulus materials and had identified the anaphors as such on the basis of recognition memory capacities. If this had been the case, then the re-appearance in the current passage of a noun that had previously appeared in the passage would have constituted a clue to the effect that a noun was a repeated-noun anaphor. Participants would have been alert, at some level, for the occurrence of such clues. Because sentences functioned as passages in this experiment, participants would have been alert while reading each sentence, for the occurrence of nouns that had previously appeared in that sentence. When a participant recognized a noun as having appeared previously in the current sentence, other words in the sentence would have been subject to a recognition decrement. Words from the previous sentence would not, however, have been subject to a recognition decrement. The key difference between the One and the Two Sentence conditions was that the second test word came from the current sentence in the former condition and from the previous sentence in the latter condition. In the Two Sentence condition, therefore, the second test word should not have been subject to a recognition decrement when the participant recognized the anaphoric noun in the second sentence as having appeared previously in that sentence.

At the same time, these results are difficult to accommodate under the anaphor irrelevant interpretation of the results of Experiments 1A and 1B. They are not what would have been expected if participants had ignored the discourse properties of the stimulus materials. If this had been the case,

then, in order to account for the recognition decrement that occurred in the One Sentence condition, we would have to assume that participants were sensitive to repeated words in the stimulus items, on the basis of some automatic recognition process. We would have to assume that there was a memory set associated with this automatic recognition process and that other words in this memory set, including the second test word, were subject to a recognition decrement following the recognition of the anaphoric noun, through a process such as Dopkins and Ngo have proposed. In order to account, at the same time, for the fact that a recognition decrement was not present in the Two Sentence condition, we would have to stipulate that participants were only sensitive to the repetition of words within the context of the current sentence, and thus that the memory set associated with recognition process contained the words from the current but not the previous sentence.

This account is not very convincing. First, it is somewhat implausible that participants would only be sensitive to the repetition of words within the context of *current sentence* given that participants were required to respond positively to recognition probes if they appeared during the *current trial*. It is more plausible that participants would have limited their repetition monitoring to the current sentence under the anaphor relevant interpretation. Under this account, participants were monitoring to detect anaphors, and the repetition of a word was a sign that the word might be an anaphor only if the repetition occurred in the current sentence. Second, it is not clear how participants could have limited their repetition monitoring to the current sentence given that they were ignoring the discourse properties of the stimulus materials. In short, the present results are difficult to accommodate under the anaphor irrelevant interpretation.

Experiment 3

As has been noted, the explanation that the anaphor irrelevant account provides for the results of Experiment 2 is somewhat unconvincing. Experiment 3 sought to argue even more conclusively against this explanation by explicitly inducing participants to behave in the manner envisioned under the

explanation and showing that the pattern of Experiment 2 did not occur. The experiment was basically a replication of the One and Two Sentence conditions of Experiment 2, with the words in the stimulus sentences being scrambled such that only two of the words continued to hold their original positions, the words in question being a) the word that was later presented as the second test word and b) the anaphor that ended the only sentence in the One Sentence/After condition and the second sentence in the Two Sentence/After condition. On each trial, the participant read through a stimulus item and responded to two recognition probes, indicating for each, whether the test word had appeared *during the current trial*. After responding to the second probe, participants indicated whether any words had been repeated *in the current scrambled sentence*.

The results of interest were response time and error rate for the second test word. As in Experiment 2, the key to the experimental predictions was the last word processed before the second recognition probe in the After condition. Whereas this *item-ending* word was an anaphor in Experiment 2, in the present experiment it was simply a noun that was repeated from earlier in the current scrambled sentence. The experiment asked whether a recognition decrement would occur for the second test word when participants processed the item-ending word. The object was to test for a recognition decrement in the One and Two Sentence conditions.

According to the anaphor irrelevant explanation of the results of Experiment 2, a recognition decrement should occur in the One but not the Two Sentence condition. Participants will be induced to behave in the way that they are assumed to have behaved in Experiment 2. That is, participants will ignore the discourse properties of the stimulus materials, because the materials are scrambled, and participants will be sensitive to the repetition of a word in the current scrambled sentence, because the repeated-word question explicitly requires this. Given that participants are induced to behave in this way, the pattern of Experiment 2 should be observed; a recognition decrement should occur in the One Sentence but not the Two Sentence condition.

In more detail, the anaphor irrelevant explanation predicted as follows: When participants recognize the item-ending word as having appeared before in the current scrambled sentence, other words in that scrambled sentence will be subject to a recognition decrement, through a process such as Dopkins and Ngo (2002) proposed. Words in the previous scrambled sentence will not be subject to such a recognition decrement because they will not be contained in the memory set associated with the recognition process. In the One Sentence condition, the two pieces of text will constitute a single scrambled sentence. Words in that scrambled sentence, including the second test word, will be contained in the memory set and will be subject to a recognition decrement. In the Two Sentence condition, the two pieces of text will constitute two scrambled sentences. Words in the first scrambled sentence, including the second test word, will not be contained in the memory set and will not be subject to a recognition decrement.

Method

Participants. The participants were 24 students at the George Washington University. They received extra credit in a psychology course in exchange for their efforts. Four participants were replaced because their overall error rate for responses to the second test word was greater than .5.

Design. Number of Sentences (One, Two) and Probe Position (Before, After) were manipulated within participants and items.

Materials. The 28 experimental items were derived from the items of the One Sentence and Two Sentence conditions of Experiment 2, as follows: 1) Recall that, in the items for Experiment 2, the single sentence in the One Sentence condition typically contained 2 anaphoric relations whereas the second sentence in the Two Sentence condition contained a single anaphoric relation. If the items had been used without alteration, this inequity would have been problematic for interpretation of the repeated-word data; at the point that the repeated word question occurred, the current scrambled sentence would have contained two repeated words in the One Sentence condition and a single

repeated word in the Two Sentence condition. The items were changed therefore, so that the single sentence in the One Sentence condition contained a single anaphoric relation. 2) The words that comprised each item were scrambled according to a different random pattern for each participant. In the scrambling process, a word remained within the boundaries of its native piece of text, and two words – a) the word later presented as the second test word and b) the item-ending word - continued to hold their original positions. 3) The first word in each scrambled sentence was capitalized and the last word was followed by a period.

The experimental items were randomly intermixed with 42 filler items. For four of the filler items, the stimulus material consisted of a single scrambled sentence and the correct response to both test words was "yes." Of these items, the test words were both nouns for one item, a verb and noun for one item, and both verbs for two items. For five of the filler items, the stimulus material consisted of a single scrambled sentence and the correct responses to the test words were "yes" and "no" respectively. Of these items, the test words were both nouns for one item, a noun and verb for three items, and both verbs for one item. For thirteen of the filler items, the stimulus material consisted of a single scrambled sentence and the correct response to both test words was "no." Of these items, the test words were both nouns for five items, a noun and verb for four items, and both verbs for four items. For one of the filler items, the stimulus material consisted of two scrambled sentences and the correct response to both test words was "yes". For this item, the test words were a verb and a noun. For four of the filler items, the stimulus material consisted of two scrambled sentences and the correct responses to the test words were "yes" and "no" respectively. Of these items, the test words were both nouns for three items, and a verb and a noun for one item. For fifteen of the filler items, the stimulus material consisted of two scrambled sentences and the correct response to both test words was "no". Of these items, the test words were both nouns for six items, a verb and a noun for seven items, and both verbs for two items.

In twenty-two of the one-sentence filler items, the same noun occurred twice. For these items,

the correct response to the repeated-word question was “yes.” In twelve of the two-sentence filler items, the same noun occurred in the first and second scrambled sentence. For these items, the correct response to the repeated-word question was “no.” In eight of the filler items, no word occurred more than once. For these items, the correct response to the repeated-word question was again “no.”

Procedure. Each participant was randomly assigned to a presentation set. The experiment was presented on a computer monitor, according to a procedure that followed that of the earlier experiments in all respects except that the presentation time for each word in the list was 30% longer. The participant was instructed to indicate, for each recognition probe, whether the test word had appeared during the current trial. After the participant responded to the second probe, the test word disappeared and a prompt appeared at the bottom of the screen, asking whether any words had been repeated in the current scrambled sentence. Participants were instructed that each scrambled sentence would begin with a capital letter and end with a period.

Results and discussion

Performance on the first recognition probe was good and did not differ as a function of the experimental conditions (See Table 3). Response time: Number of Sentences, $F1(1,20) < 1$, $F2(1,24) < 1$; Probe position, $F1(1,20) < 1$, $F2(1,24) < 1$; Number of Sentences X Probe Position, $F1(1,20) < 1$, , $F2(1,24) < 1$. Error rate: Number of Sentences, $F1(1,20) < 1$, $F2(1,24) < 1$; Probe position, $F1(1,20) < 1$, $F2(1,24) < 1$; Number of Sentences X Probe Position, $F1(1,20) = 1.42$, $MSe = .007$, $F2(1,24) = 1.42$, $MSe = .009$.

Insert Table 3 about here

Performance on the second recognition probe was of primary interest. Because the goal of the experiment was to observe how recognition judgments to the second probe were affected as a consequence of positive recognition judgments to the item-ending word, analysis was restricted to trials on which the repeated-word question was answered correctly. The results of this analysis are

summarized in Table 3. Response time did not differ in the One and Two Sentence conditions, $F1(1,20) = 1.66$, $MSe = 89{,}991$, $F2(1,24) = 3.35$, $MSe = 252{,}083$, or the Before and After conditions, $F1(1,20) < 1$, $F2(1,24) < 1$. The effects of Number of Sentences and Probe Position did not interact in the response-time data, $F1(1,20) < 1$, $F2(1,24) < 1$.

The error rate was greater in the Two than the One Sentence condition, $F1(1,20) = 11.28$, $MSe = .036$, $F2(1,24) = 15.59$, $MSe = .040$. The error rate was also greater in the After than the Before condition, $F1(1,20) = 14.16$, $MSe = .039$, $F2(1,24) = 11.78$, $MSe = .060$. The effects of Number of Sentences and Probe Position did not interact in the error-rate data, $F1(1,20) < 1$; $F2(1,24) < 1$..

The results for the repeated word question are presented in Table 4. Performance was more accurate in the After than the Before condition, $F1(1,20) = 5.17$, $MSe = .07$; $F2(1,24) = 9.28$, $MSe = .048$, but did not differ in the One and Two Sentence conditions, $F1(1,20) < 1$; $F2(1,24) < 1$. The effects of Number of Sentences and Probe Position did not interact, $F1(1,20) < 1$; $F2(1,24) < 1$. The basis for the Before/After difference is not clear. The important result for other aspects of the experiment is that the results did not differ as a function of number of sentences.

Insert Table 4 about here

A recognition decrement occurred for the second test word in both the One and the Two Sentence conditions. In both conditions, the second test word was recognized less well in the After than the Before condition. As in Experiment 1A and Experiment 2, the recognition decrement showed up in the error-rate but not the response-time data. Again, a speed-accuracy-trade-off did not occur.

The results for the second probe are not consistent with the anaphor irrelevant explanation of the results of Experiment 2. In the present experiment, participants were explicitly induced to behave in the way this explanation says they behaved in Experiment 2. The present experiment did not produce the same pattern of results as Experiment 2; in Experiment 2, a recognition decrement occurred in the

One but not the Two Sentence condition; in the present experiment, a recognition decrement occurred in both the One and Two Sentence conditions. The failure of the present experiment to produce the pattern of Experiment 2 undermines the anaphor irrelevant explanation of the results of that experiment.

These results argue also against the anaphor irrelevant explanation in a slightly different way. To deal with the recognition probe in Experiment 3, participants had to maintain a memory set containing the words from the current trial. To deal with the repeated-word question, participants had to maintain a memory set containing the words from the current scrambled sentence. In the One Sentence condition, the two sets should have been one and the same, because a single scrambled sentence was presented on a given trial. In the Two Sentence condition, the sets should have been distinct, because two scrambled sentences were presented on a given trial. Participants evidently had trouble maintaining distinct sentence and trial sets in the Two Sentence condition. We infer this from the fact that a recognition decrement occurred for the second test word following the processing of the item-ending word. When the item-ending word was recognized as having been repeated, other words in the sentence set should have become more difficult to recognize, through a process such as Dopkins and Ngo (2002) proposed. The second test word should not have been contained in the sentence set, given that this word did not appear in the second scrambled sentence. The second test word should have been contained in the trial set, however. From the fact that a recognition decrement occurred for the second test word, we infer that the sentence and trial sets were not distinct.

Although participants in Experiment 3 had trouble maintaining distinct sentence and trial sets, participants in Experiment 2 evidently did so successfully. Like participants in Experiment 3, participants in Experiment 2 had to maintain a trial set to deal with the recognition probe. By assumption, participants in Experiment 2 also maintained a sentence set to support recognition of potential anaphors. As in Experiment 3, the two sets should have been distinct in the Two Sentence condition. Participants were evidently successful at maintaining the distinction between the sets. We infer this from the fact that,

in contrast to the case of Experiment 3, a recognition decrement did not occur for the second test word in the Two Sentence condition of Experiment 2.

Thus, whereas participants could maintain distinct sentence and trial sets in Experiment 2, when the stimulus materials were intact, they could not do this in Experiment 3, when the stimulus materials were scrambled. Evidently, discourse properties of the materials for Experiment 2 helped participants maintain the distinct memory sets. This is inconsistent with the anaphor irrelevant explanation, which says that participants ignored the discourse properties of the materials for Experiment 2.

General Discussion

The results of the present experiments have implications for our understanding of the results that Gernsbacher (1989) and Dopkins and Ngo (2002) reported. Gernsbacher (1989) observed that a noun from an intact sentence was subject to a recognition decrement following the processing of a repeated-noun anaphor in the sentence. Dopkins and Ngo observed that a noun from a scrambled sentence was subject to a recognition decrement following the processing of a noun that was repeated from earlier in the sentence. In the present study, two possibilities were considered regarding the relationship between the Gernsbacher and Dopkins and Ngo phenomena. On one hand, the two phenomena might reflect distinct processes, with the Gernsbacher phenomenon reflecting an anaphor comprehension process such as Gernsbacher has proposed, and the Dopkins and Ngo phenomenon reflecting a memory process such as Dopkins and Ngo have proposed. On the other hand, the two phenomena might reflect the same memory process.

The results of Experiments 1A and 1B suggest that the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena reflect the same memory process. In these experiments, a noun that was part of the experience of processing an intact passage, without actually being part of the passage, was subject to a recognition decrement following the processing of a repeated-noun anaphor in the passage. These results are inconsistent with the view that the Gernsbacher phenomenon reflects an anaphor

comprehension process. If the Gernsbacher phenomenon reflected such a process, then its purpose would be to suppress competitors of current anaphor's antecedent. The process should have no impact on words that are not part of the current passage and therefore not competitors of the anaphor's antecedent.

At the same time, the results of Experiments 1A and 1B are consistent with the view that the Gernsbacher (1989) and Dopkins and Ngo (2002) phenomena both reflect a memory process. Under this interpretation, a recognition decrement occurs when an anaphoric noun is recognized as having appeared before in a passage; other words in the memory set associated with the recognition of the noun become more difficult to recognize at this point. It would be reasonable under this interpretation for words that were part of the experience of comprehending a passage to be included in the recognition memory set. Thus, it would be reasonable for such words to be subject to a recognition decrement following the recognition of the anaphoric noun.

Given the conclusion implied by the results of Experiments 1A and 1B, what are the implications for our understanding of anaphor comprehension? Two possibilities were considered in the present study. According to one view, the findings of Experiments 1A and 1B have implications for our understanding of anaphor comprehension. The identification of a repeated-noun anaphor as such depends on memory capacities. An initial clue that a noun is a repeated-noun anaphor emerges by virtue of the fact that the noun is recognized as having occurred before in the current passage.

According to the other view, the findings of Experiments 1A and 1B have implications for current practice in the study of anaphor comprehension but not for our understanding of anaphor comprehension. Participants in recognition probe tasks ignore the discourse properties of the stimulus materials and treat them as lists of words upon which memory tests are to be given. Previous results obtained with such tasks are consequently called into question.

The results of Experiments 2 and 3 argue in favor of the former interpretation and against the

latter. In Experiment 2, a noun from a previous sentence was not subject to a recognition decrement following the processing of a repeated-noun anaphor in the current sentence. These results argue against the anaphor irrelevant interpretation of the results of Experiments 1A and 1B. They are inconsistent with the view that participants in the Gernsbacher (1989) task ignore the discourse properties of the stimulus materials. If participants had ignored the discourse properties of the materials for Experiment 2, then they should not have been able to tell the difference between the current and the previous sentence.

At the same time, the results of Experiment 2 argue for the anaphor relevant interpretation of the results of Experiments 1A and 1B. They are consistent with the view that the identification of a repeated-noun anaphor as such depends on memory capacities. If this is the case, participants should have been sensitive, at some level, while reading each sentence in the Two Sentence condition, to the occurrence of words that had already appeared in that sentence. When participants recognized the anaphoric noun as having appeared before, other words in the current sentence should have been subject to a recognition decrement. Words from the previous sentence, including the test word, should not have been subject to such a decrement.

In Experiment 3, which was similar to Experiment 2 except that the stimulus materials were scrambled, a noun from a previous sentence was subject to a recognition decrement following the processing of a word that had been a repeated-noun anaphor in the unscrambled version of the current sentence. These results argue against the anaphor irrelevant interpretation of the results of Experiments 1A and 1B. If participants in the Gernsbacher (1989) task ignore the discourse properties of the stimulus materials, then similar outcomes should have been observed in Experiments 2 and 3.

The present results imply, then, that recognition memory capacities play a role in anaphor comprehension. Before an anaphor can be resolved, it must be identified as an anaphor. Pronouns have superficial properties that mark them as potential anaphors but nouns do not. The definite article can be used to mark nouns as potential anaphors, but such markings are not always present. The present

results suggest the existence of a completely different marking mechanism: Nouns may be identified as potential anaphors by virtue of the fact that they are recognized as having occurred before in the current passage.

Such a marking mechanism would clearly be effective in the case of repeated-noun anaphors. It might also be effective in the case of other noun anaphors (e.g. 'guy' as an anaphor with antecedent 'salesman'). In the latter case, the mechanism would have to indicate the previous occurrence, not of a particular word, but rather of a word belonging to the category indicated by a particular word.

We infer the existence of the proposed anaphor marking mechanism from evidence that recognition decrements with intact and scrambled passages reflect the same underlying process. It must be acknowledged that the intact passages in question have been somewhat less than naturalistic. Most crucially, the passages have been rather short. With such short passages, explicit anaphor markers like the definite article may be of reduced effectiveness. This follows because one does not violate convention in a short passage when one introduces a character for the first time with the definite article. Thus, it is possible that the recognition-based marking mechanism plays a larger role with such short passages than with longer passages.

The present results imply that a recognition decrement occurs for the words in a passage whenever a repeated-noun anaphor is processed. What is the functional significance of this recognition decrement for comprehension of the anaphor? Notice that the question of the functional significance of the recognition decrement is independent of the question of the functional significance of recognition memory. Regardless of the recognition decrement's functional significance, it signals a role for recognition memory in anaphor processing. What then is the significance of the recognition decrement? On one hand, it may index an accessibility decrease that serves a useful function for comprehension. For example, such an accessibility decrease might facilitate anaphor resolution in the manner of Gernsbacher's proposed process. Notice, though, that in this case, the facilitation of anaphor resolution

would be a windfall from a memory process rather than the direct product of a comprehension process. On the other hand, the recognition decrement may have no functional significance. It may simply be a byproduct of the anaphoric noun's recognition. We will be able to better address these questions when we better understand the processes underlying the recognition decrement.

It must be noted that recognition decrements have been observed following the processing of pronoun as well as repeated-noun anaphors (Gernsbacher, 1989; McDonald & MacWhinney, 1990). Pronoun decrements may reflect a somewhat different process than repeated-noun decrements do. In support of this idea, Gernsbacher (1989) observed that, whereas repeated-noun decrements occurred immediately after the processing of repeated-noun anaphors, pronoun decrements occurred more subsequently following the processing of pronouns. Granting that repeated-noun decrements reflect memory processes, we must consider the possibility that pronoun decrements also reflect such processes. In the case of pronouns, the process would be associated with recall rather than recognition. In line with Gernsbacher's results, recall of the pronoun's antecedent would occur at some remove following that processing of the pronoun.

At the same time that they imply a role for recognition memory in anaphor comprehension, the present results offer a vote of confidence for the recognition probe task as a means of studying comprehension. Results obtained with the task have been called into question previously on the grounds that participants ignore the discourse properties of the stimulus materials and treat them as lists of words on which memory tests are to be given (Gordon, et al., 2000). The results of Experiments 2 and 3 imply that, at least with the recognition probe task of the present study, this was not the case.

What are we to make of the present recognition decrement as a memory phenomenon? It bears an intriguing relationship to another phenomenon associated with the recognition of words from discourse. Previously it has been shown that a word from a sentence is subject to a recognition increment when the word is tested following the recognition of a word from the same sentence. The

increment is greatest when the two words belong to the same proposition (Ratcliff & McKoon, 1978, see also McKoon & Ratcliff, 1980). Why is recognition facilitated in this previous case and impeded in the present case? The answer probably lies in details of the procedures under which the two phenomena have been observed.

A recognition increment has been observed when participants have made a positive recognition judgment to a word from a sentence after having previously made a positive recognition judgment to a word from the same as opposed to a different sentence. The relevant comparison in the present study contrasts performance on the second recognition probe in the One Sentence/After and Two Sentence/After conditions. On the basis of past demonstrations of recognition increment, we would expect performance to be better in the One Sentence/After than the Two Sentence/After condition. In fact, we find an advantage for the One Sentence condition in the error rate data, $F1(1,36) = 3.61$, $MSe = .025$ (one-tailed), $F2(1,24) = 4.79$, $MSe = .013$. (Note that we also find a nominal but non-significant advantage for the Two Sentence condition in the response-time data, $F1(1,36) = 1.91$, $MSe = 96{,}698$, $F2(1,24) < 1$).

In the present study, a recognition decrement was observed when participants made a positive recognition judgment to a word from a passage after deciding that another word from the same passage either did or did not occur previously in the passage. In other words, participants made a positive recognition judgment after making a positive as opposed to a negative recognition judgment to another word from same passage. A recognition decrement may not have had a chance to display itself in the studies that have observed a recognition increment because the relevant comparison between positive and negative judgment was not posed.

Further work will be necessary to determine the processes underlying the present recognition decrement. Among other things, it will be important to determine whether the decrement involves a decrease in sensitivity or a shift in response bias. The answer to this question has no ramifications for the

most important implication of the present results - that recognition memory plays a role in the identification of anaphors. The recognition decrement marks the involvement of recognition memory in anaphor comprehension equally well regardless of whether the decrement involves a change in sensitivity or bias.

The answer to sensitivity/bias question may have implications regarding the functional significance of the recognition decrement for anaphor comprehension. If the recognition decrement involves a change in sensitivity, then we can easily see it as reflecting a process that reduces the accessibility of competitors to an anaphor's antecedent. One would expect a reduction in accessibility to show up both as a reduction in recognition sensitivity and as a reduction in availability to recall given an anaphoric cue. On the other hand, if the recognition decrement involves a change in bias, then we may have a harder time seeing the decrement as reflecting an accessibility reduction process. It is not clear that a relationship should exist between a change in recognition criterion and a change in recall availability.

In conclusion, the present results suggest that the same memory process underlies the recognition decrement that occurs when a repeated-noun anaphor is processed in an intact sentence and the recognition decrement that occurs when a repeated noun is processed in a scrambled sentence. Furthermore, the results suggest that this is not because participants in recognition probe tasks ignore the discourse properties of the stimulus items and treat them as simple lists of words upon which memory tests are to be given.

References

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, ***106***, 748-765.

Cacciari, C., Carreiras, M., & Cionini, C. B. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory and Language*, ***37***, 517-532.

Cloitre, M., & Bever, T. G. (1989). Linguistic anaphors, levels of representation, and discourse. *Language & Cognitive Processes*, ***3***, 293-322.

Corbett, A. T., & Chang, F. R. (1983). Pronoun disambiguation: Accessing potential antecedents. *Memory & Cognition*, ***11***, 283-294.

Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior*, ***22***, 121-132.

Dopkins, S., & Ngo, C. T. (2002). Inhibition of verbal memory retrieval as a consequence of prior retrieval. *Journal of Memory and Language*, ***46***, *606-621*.

Dopkins, S., & Ngo, C. T. (2004). Exploring a recognition-induced recognition decrement. Submitted.

Garnham, A. (1985). *Psycholinguistics*. London: Methuen.

Garnham, A. (1987). Understanding anaphora. In A. W. Ellis (Ed.), *Progress in the Psychology of Language, (vol. 3)*. London: Erlbaum.

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. London: Methuen.

Garnham, A. Oakhill, J., & Cain, K. (1997). The interpretation of anaphoric phrases: Time course, and effects of overspecificity. *Quarterly Journal of Experimental Psychology*, ***50A***, 149-162.

Garrod, S., Freudenthal, D., & Boyle, E. A. (1994). The role of different types of anaphor in

the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, ***33***, 39-68.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, ***32***, 99-156.

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

Glenberg, A. M., & Kruley, P. (1992). Pictures and anaphora: Evidence for independent processes. *Memory & Cognition*, ***20***, 461-471.

Gordon, P. C., Hendrick, R., & Foster. K. L. (2000). Language comprehension and probe-list memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, ***26***, 766-775.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, ***27***, 1411-1423.

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, ***13***, 425-430.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, ***101***, 371-395

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, ***99***, 122-149.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review*, ***95***, 163-182.

MacDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, ***34***, 543-566.

McDonald, M. C., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. *Journal of Memory and Language*, ***29***, 469-492.

McKoon, G., & 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.

Nordlie, J. W., Dopkins, S., & Johnson, M. (2001). Words in a sentence become less accessible when an anaphor is resolved. *Memory & Cognition*, ***2***, 201-220.

O'Brien, E. J. (1987). Antecedent search processes and the structure of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, ***13***, 278-290.

O'Brien, E. J., Duffy, S. A., & Myers, J. L. (1986). Anaphoric inference during reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, ***12***, 346-352.

O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. (1990). Antecedent retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, ***16***, 241-249.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior*, ***17***, 403-417.

Table 1

Sample stimuli

Experiments 1A and 1B

Before condition

Linda thought that Agnes would never return alive from her shopping trip. Debra had asked SUN for a rich chocolate cake but everyone thought that was a bad idea and SUN

After condition

Linda thought that Agnes would never return alive from her shopping trip. Debra had asked SUN for a rich chocolate cake but everyone thought that was a bad idea and Linda SUN

Experiment 2

One Sentence/Before condition

The proprietor of the apartment ejected the artist who had painted the star AGENT so the artist had to send ten chickens to the STAR

One Sentence/After condition

The proprietor of the apartment ejected the artist who had painted the star AGENT so the artist had to send ten chickens to the proprietor STAR

Two Sentence/Before condition

The proprietor of the apartment ejected the artist who had painted the star AGENT The farmer punished the cow that had stepped on the STAR

Two Sentence/After condition

The proprietor of the apartment ejected the artist who had painted the star AGENT The farmer punished the cow that had stepped on the farmer STAR

Replication/Before condition

The mechanic warned the professor about the car the student was selling AGENT but the professor had owned an old car for years and ignored the AGENT

Replication/After condition

The mechanic warned the professor about the car the student was selling AGENT but the professor had owned an old car for years and ignored the mechanic AGENT

Note. The upper case word was presented at the indicated points as a recognition probe.

Table 2

Results of Experiment 2

Probe position		Materials type		
		Replication	One	Two
		First probe		
Before	Response time	1106 (50)	1092 (56)	1174 (64)
	Error rate	.07 (.02)	.07 (.02)	.06 (.02)
After	Response time	1163 (63)	1143 (71)	1131 (61)
	Error rate	.05 (.01)	.08 (.02)	.05 (.02)
		Second probe		
Before	Response time	1472 (93)	1427 (70)	1439 (69)
	Error rate	.36 (.04)	.15 (.02)	.31 (.04)
After	Response time	1416 (81)	1410 (90)	1314 (64)
	Error rate	.50 (.04)	.29 (.04)	.36 (.04)

Note. The standard error of the mean is given in parentheses.

Table 3

Experiment 3: Recognition probe results

Probe position		Materials Type	
		One	Two
		First probe	
Before	Response time	1022 (44)	975 (52)
	Error rate	.05 (.02)	.04 (.02)
After	Response time	998 (46)	986 (58)
	Error rate	.06 (.01)	.06 (.02)
		Second probe	
Before	Response time	1426 (114)	1514 (139)
	Error rate	.17 (.04)	.32 (.04)
After	Response time	1487 (167)	1561 (159)
	Error rate	.33 (.04)	.44 (.05)

Note. The standard error of the mean is given in parentheses.

Table 4

Experiment 3: Error rate on repeated word question

Probe position	Materials type	
	One	Two
Before	.33 (.04)	.36 (.05)
After	.21 (.04)	.21 (.03)

Note. The standard error of the mean is given in parentheses.

Appendix

Sample experimental items from Experiments 1A and 1B

Anne had heard Jane give a frank analysis in the locker-room. They both MONDAY expected that Pam would lose the track meet but Anne MONDAY
Did Jane and Anne talk in the bathroom?
Did Anne expect that Pam would lose?

George found Sam hanging by one arm from a tree branch. Then Ray MOUTH called to ask if anyone wanted to go swimming and George MOUTH
Did Ray want to go swimming?
Did Sam find George in the pool?

Linda thought that Agnes would never return alive from her shopping trip. Debra had asked SUN for a rich chocolate cake but everyone thought that was a bad idea and Linda SUN
Did Agnes go on a shopping trip?
Was Debra going to make a pound cake?

Jack had seen the note that had Bob left by the restroom. Soon after, Billy SHOE offered to sell a car that everyone thought was jinxed and Jack SHOE
Had Jack seen the note?
Does Bob drive a jinxed car?

Phil departed from the cabin at the lake after Frank. Dick was sound DINNER asleep and nobody wanted to wake him so Phil DINNER
Did Frank leave before Phil?
Was Dick eating in the cabin?

Sample filler items from Experiment 1A and Experiment 1B

In order to make a statue of the ARTIST professor, Joel found Kent some marble and some sculpting tools and Joel FOUND
Did Joel find the sculpting tools?
Was the statue going to be made of marble?

After visiting Patty, the accountant sent Becky a check for SENT ten grand and Becky ACCOUNTANT
Did the accountant send a check?
Was there a reason stated for sending the check?

In the eyes of the saint, Mark had done Jack EYES a great wrong by BIBLE
Did the saint feel that Mark had sinned?
Was the name of the saint revealed?

The artist asked his collaborator ARTIST to paint the other half of the canvas and OIL
Was the artist making a sculpture?

Was the collaborator supposed to paint more than half?

As the dignitary emerged from his stretch limo, DIPLOMAT Amy handed Celia the camera and LIMO
Did Amy hand Celia a revolver?
Were Amy and Celia in the stretch limo?

Sample experimental items from One and Two Sentence conditions of Experiment 2 (One sentence condition followed by Two sentence condition)

The waiter showed the bartender a gun after he slapped the busboy KILLER but the bartender was nearsighted and didn't understand the waiter BUSBOY
Did the bartender slap the busboy?

The waiter showed the bartender a gun after he slapped the busboy KILLER The teacher spoke with the student who hated the teacher BUSBOY
Did the student hate the teacher?

The hustler asked the go-between to challenge the detective PIMP but the go-between didn't hear the hustler DETECTIVE
Did the detective challenge the hustler?

The hustler asked the go-between to challenge the detective PIMP The liontamer thought highly of the liontamer DETECTIVE
Did the detective challenge the hustler?

The manager predicted that the pitcher would walk the slugger CATCHER but the pitcher just laughed at the manager SLUGGER
Did the manager predict that the slugger would walk?

The manager predicted that the pitcher would walk the slugger CATCHER The president fulfilled his responsibilities as the president SLUGGER
Did the manager predict that the slugger would walk?

The chef gave the reporter the sweater the actress had left FAN and the reporter made a cake for the chef ACTRESS
Did the reporter make a cake?

The chef gave the reporter the sweater the actress had left FAN The carpenter made a beautiful coffin for the carpenter ACTRESS
Did the chef give away a sweater?

The realtor got the buyer to sleep on the sofa the owner had left HANDYMAN but the buyer had a bad dream and blamed the realtor OWNER
Did the buyer go to sleep?

The realtor got the buyer to sleep on the sofa the owner had left HANDYMAN The king fired the advisor who had plotted against the king OWNER

Did the buyer go to sleep?

Sample experimental items from Replication condition of Experiment 2

After the press conference with the senator, the lobbyist played golf with his opponent JUDGE after which the lobbyist went out to dinner with the senator JUDGE
Did the senator play golf?

The detective left the suspect alone with the victim CORPSE and the suspect searched the room for clues to protect himself against the detective CORPSE
Did the detective leave the suspect alone?

After the misleading encounter with the drifter the fisherman got help and directions from the ranger TRAVELER but the fisherman never forgave the drifter TRAVELER
Did the fisherman ask the ranger for directions?

The bartender tipped off the marshal concerning the arrival of the killer VICTIM so the killer had to exact revenge from the bartender VICTIM

In the sauna the mobster handed the bimbo a strong drink for the senator REPORTER but the senator was wise to the ways of the mobster REPORTER
Did the senator have a drink?

Sample filler items from Experiment 2

The singer broke down and cried to the priest in the slum BUM when he heard that the actress was gravely ill DIED
Was the singer gravely ill?

The administrator spilled a drink on the professor at the holiday party SPILLED so the professor refused to turn in his grades REFUSED
Did the professor spill his drink?

The guide offered to take a photo of the famous writer at the waterfall POET The little lost girl was found by the guard near the lion cage MOTHER
Was the girl found by the lion?

The butcher counted the money that the chef had left GROCER while the chef looked frantically for the therapist in the parking lot FOUND
Did the chef look for someone in the parking lot?

The conductor called out the station as the engineer braked the train WHISTLED then the conductor pinched the scientist and tried to wake him COMMUTER
Did the engineer pinch the scientist?