Journal of Experimental Psychology:
Learning, Memory, and Cognition
2000, Vol. 26, No. 3, 766–775

0278-7393/00/$5.00 DOI: 10.1037//0278-7393.26.3.766

Language Comprehension and Probe-List Memory

Peter C. Gordon, Randall Hendrick, and Kerry Ledoux Foster
University of North Carolina at Chapel Hill

Experiments were performed using probe-word recognition methodology in which participants read sentences that were presented 1 word at a time and were then shown a probe word and had to make a speeded response indicating whether the word had occurred in the sentence. One experiment showed that response times to probe words increased with the size of the set of candidate probes. The other experiments showed that the effects caused by name repetition in circumstances in which the repeated name was co-referential also occurred when the repeated name was not co-referential and when the order of words in a sentence was scrambled. The results suggest that responses in the task can be based on *probe-list memory,* a mental representation created to keep track of those words that the participant believes are likely to be probed, and that the use of the task to make inferences about language comprehension should be accompanied by controls ruling out such strategies.

Probe-word recognition tasks have been used extensively in experiments on language comprehension, and the results of those experiments have played an important role in supporting claims about the general nature of language processing, as well as claims about specific components of language comprehension (e.g., Gernsbacher, 1989; Greene, McKoon, & Ratcliff, 1992). In a probe-word recognition task, a participant reads or hears a sentence (or other fragment of language). At some point, a probe word is presented, and the participant must respond as rapidly as possible as to whether or not the probe word occurred in the sentence. The response time to the probe is used to make inferences about the accessibility of concepts in the memory representation of the sentence. The probe-word task is very flexible. It can be used with auditory or visual presentation of the linguistic material (Caplan, 1972), processing time can be controlled by manipulating the speed of presentation of the linguistic material (Greene et al., 1992), comprehension strategies can be influenced by the focus of comprehension questions (Greene et al., 1992), and the time course of language comprehension can be charted by varying the timing (or position) of the probe word (Gernsbacher, 1989). The task has been used to measure a number of aspects of language comprehension, including the syntactic organization of the memory representation of a sentence (Caplan, 1972), the propositional (semantic) organization of a memory representation (McKoon & Ratcliff, 1980), the extent of inference in language comprehension (McKoon & Ratcliff, 1992), and the mechanisms by which reference and co-reference are processed (Chang, 1980). The research in this article examined a critical assumption underlying the use of the probe-word task to make inferences about language comprehension, particularly as it has been used in studies of reference and co-reference.

The assumption examined in this article is that response times to the probe word reflect the accessibility (or activation) of a word (or concept) in the memory representation of the sentence (or other language fragment) that is being understood. This assumption is critical in using results of the probe-word task to make inferences about language comprehension; we call it the *sentence-memory assumption.* Other assumptions about this task have been investigated by determining how the relative frequency of different types of trials affects patterns of response times (Ratcliff & McKoon, 1981), by analyzing the appropriate baselines for comparing response times (MacDonald & MacWhinney, 1990), and by examining strategic effects on pronoun interpretation (Greene et al., 1992). However, to our knowledge, the sentence-memory assumption has received only little scrutiny. In particular, Dell, Ratcliff, and McKoon (1981) showed that repeated testing with the same probe words could lead to responses based on memories of prior test episodes rather than on the sentences that had just been processed for comprehension purposes. Other than that study, the sentence-memory assumption appears to have been accepted without examination.

As an alternative to the sentence-memory assumption, it is possible that participants in the probe-word recognition task engage in a strategic adaptation to the dual-task demands of language comprehension and responding in the probe-word task. More specifically, in addition to creating memory representations for the sentence, participants may also be creating memory representations for use in respond-

Peter C. Gordon, Randall Hendrick, and Kerry Ledoux Foster, Department of Psychology, University of North Carolina at Chapel Hill.

This research was supported by National Science Foundation Grant SBR-9807028.

We thank Elliot Hirshman and the *Journal of Experimental Psychology: Learning, Memory, and Cognition* reviewers for helpful comments on an earlier version of this article. We also thank Chin Lung Yang, Michael Boyd, Susan Carawan, Marcus Johnson, and Sara Tate for help in conducting the research.

Correspondence concerning this article should be addressed to Peter C. Gordon, Department of Psychology, University of North Carolina, Chapel Hill, North Carolina 27599-3270. Electronic mail may be sent to pcg@email.unc.edu.

ing in the probe task, a possibility that we call the *probe-list hypothesis*. The sentence-memory hypothesis leads to the prediction that performance in the probe task should be influenced only by the memory processes involved in understanding the passage, whereas the probe-list hypothesis leads to the prediction that memory demands imposed by the probe task itself should also affect performance in the probe task. To the extent that the probe-list hypothesis is correct, we believe that data from probe-word recognition experiments should not be used to support theoretical accounts of language comprehension.

Our evaluation of the probe-list hypothesis was prompted by the discrepant views on co-reference that have emerged from probe-word recognition experiments and from reading-time experiments, particularly with regard to the relative effectiveness of establishing co-reference with repeated names as compared with pronouns. The results of probe-word recognition tasks have been interpreted as indicating that co-reference is established more immediately and automatically with repeated names (or with other definite expressions) than with pronouns (Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989; Greene et al., 1992; MacDonald & MacWhinney, 1990). Gernsbacher (1989, 1990, 1996), in her *Structure-Building Framework*, has argued that co-reference is more readily established with repeated names than with pronouns because repeated names provide more information than do pronouns for activating relevant memory representations. She argued that well-established principles about the operation of memory indicate that retrieval from memory is enhanced when a stimulus provides a richer set of retrieval cues, a position she called the *explicitness hypothesis*.

In contrast, research using self-paced reading-time tasks, eye tracking during reading, and judgments of coherence has shown that co-referential pronouns are often read more quickly than are co-referential repeated names (e.g., Garrod, Freudenthal, & Boyle, 1994; Gordon, Grosz, & Gilliom, 1993; Gordon & Hendrick, 1997; Hudson, Tanenhaus, & Dell, 1986; Hudson-D'Zmura, 1988; Kennison & Gordon, 1997). Evidence for greater ease in establishing co-reference with pronouns as compared with repeated definite phrases has also been obtained using cross-modal priming techniques (Cloitre & Bever, 1988). Drawing on a number of sources, we (Gordon & Hendrick, 1998) have argued that this pattern occurs because the primary function of pronouns (and other reduced expressions) is to refer to things that have already been mentioned in a discourse and that are mentally represented in a discourse model. Therefore, pronouns are the natural vehicle for co-reference. In contrast, the primary function of names (and of other full expressions) is to introduce entities into the discourse model. Therefore, co-reference with repeated names requires additional mental processes because distinct entities are introduced into the discourse model with each occurrence of the name. Our view (Gordon & Hendrick, 1998) of how co-reference is established with names and pronouns is almost directly opposite of that advanced by Gernsbacher (1989, 1990, 1996) in her *Structure-Building Framework*.

The experiments reported in this article examined whether the discrepancy between these two views may be resolved by further understanding of the probe-word recognition task. We did so by examining whether probe-list memory can contribute to responses in the probe-word recognition task under task conditions that have been used in previous research and by examining whether the effects of repeated names on response times, which have previously been attributed to the processing of co-reference, can instead be attributed to probe-list memory. Beyond issues of co-reference, the results of the experiments have implications for the position that important aspects of language comprehension can be explained by general models of memory and information processing (e.g., Gernsbacher, 1989; Greene et al., 1992). Evidence that the probe-list hypothesis is true would undermine that position by showing that some of the central results used to support it do not actually reflect language comprehension processes.

Experiment 1

Participants engaged in a strategic task adaptation would be expected to put in probe-list memory only those words that they think are likely to be probed. As the number of candidate probe words increases, then so would the length of the list of words that would be included in probe-list memory. Research on recognition memory has shown that list length has a powerful effect on the speed and accuracy of responding to probes in a number of recognition-memory paradigms (e.g., Gillund & Shiffrin, 1984; Hockley & Corballis, 1982; Sternberg, 1966; see Hockley & Murdock, 1987, for a review). As list length increases, there is an increase in response times and a decrease in accuracy both for old items that were on the list and for new items that were not on the list. This effect has been explained as being due to a decrease in the mean of the distribution of activations for old items as the number of old items on a list increases. Given a general distance-to-criterion model of response times (e.g., Hockley & Murdock, 1987), this decrease of average activation leads to the observed decrease in memory performance as list length increases. Accordingly, the probe-list hypothesis leads to the prediction that response times to probes should increase with an increase in the number of candidate words that can possibly appear as probes in the task. This experiment tested that hypothesis by comparing response times to proper-name probes for participants who were presented only with proper-name probes throughout the experimental session with response times to those same proper-name probes by other participants who were probed both with proper names and with other content words during the experimental session. If the probe-list hypothesis is correct, then response times should be longer in the condition that includes both proper-name and content-word probes than in the condition that includes only proper-name probes. In contrast, the sentence-memory hypothesis predicts that response times should not change as a function of the words that could possibly be probed because the response to the probe word is assumed to be based on the mental representation of the meaning of the sentence, which would not be influenced by the pool of candidate probe words because that has no relevance to the meaning of the sentence.

The experiment was modeled after ones presented in Gernsbacher (1989). A sample sentence is shown in Example 1, along with the probes that could possibly be used.

> (1) Bill handed John some tickets to a concert but he took the tickets back immediately.
> Possible probes in names-only condition: *BILL JOHN*
> Possible probes in mixed condition: *BILL JOHN HANDED TOOK CONCERT TICKETS*

In the names-only condition, the probe word was always a name; on positive trials, it was one of the two names that had appeared in the sentence, and on negative trials, it was a common name of the same gender as the two names that had appeared in the sentence. In the mixed condition, the probe word could be either a name or some other content word; on positive trials, it could be a name, a noun, or a verb from the first or second clause of the sentence, and on negative trials, it could be a name, a noun, or a verb that was not in the sentence. In the mixed condition, 30 of the 110 trials used name probes (half positive and half negative). The condition of names-only versus mixed was manipulated between participants, and identical instructions were given to both groups of participants so that any differences in strategies would be due to the participants' beliefs about the best way to perform the tasks.

Method

Participants. Forty-four students at the University of North Carolina at Chapel Hill served as participants in the experiment, 22 in each condition of the between-participants design. They received course credit in introductory psychology for their participation.

Stimulus materials. A total of 110 sentences were used in the experiment after the initial block of practice trials (18 sentences). Fifty-five of the sentences were taken from the appendix to Gernsbacher (1989); they served as stimuli in the positive trials in which the probe word was in the sentence, just as they had in the Gernsbacher experiments. An additional 55 sentences from other sources were adapted for this experiment; they served as the negative, probe-absent stimuli. All of the sentences in the experiment had the same general structure as those in the Example 1. They consisted of two clauses: The first clause mentioned two different named characters of the same gender, and the second clause contained a pronoun that (on semantic grounds) was co-referential with one of the two names in the first clause. In half of the sentences, the pronoun co-referred with the first-mentioned character, and in the other half, it co-referred with the second-mentioned character.

Design and procedure. The sentences were grouped into an initial warm-up block of 18 sentences followed by five experimental blocks of 22 sentences each. Every block contained an equal number of probe-present and probe-absent trials presented in a random order.

In each trial, participants read a single sentence that was presented one word at a time in the center of the screen using rapid serial visual presentation (RSVP) methodology. Each word was presented for 300 ms plus an additional 16.7 ms for each character in the word. At the end of the sentence, a single probe word in all capital letters was presented several lines above the location of the words in the sentence. Participants were required to indicate as quickly as possible whether the probe word had occurred in the preceding sentence, pressing the / key with their right index finger if it had and the Z key with their left index finger if it had not. Then, a comprehension question was presented that tested knowledge of the constituent organization of the sentence (i.e., who did what).

Results and Discussion

The results are shown in Table 1, which gives the response times for accurate responses to name probes in the mixed condition and for the matched trials using name probes in the names-only condition. Response times were significantly slower in the mixed condition than in the names-only condition both for positive responses (by 234 ms) and for negative responses (by 224 ms), $F_1(1, 42) = 7.7$, $MSE = 150{,}460$, $p < .01$, $F_2(1, 28) = 16.4$, $MSE = 47{,}942$, $p < .001$. The interaction of condition (mixed vs. names-only) with type of response (positive vs. negative) was not significant. The percentages of responses that were correct were not significantly influenced by the experimental factors.

The results of the experiment are consistent with the probe-list hypothesis. Response times both for positive and for negative responses increased as the number of candidate probe words increased; this is a finding that would be expected if performance is based on memory for a list whose length increases with the number of possible probes (Hockley & Murdock, 1987). In contrast, the sentence-memory assumption that underlies the use of the probe task to study language comprehension provides no account of the observed findings. There is no reason why the number of candidate probes should influence the memory representations that are developed as part of understanding the meaning of a sentence. Whereas the results of the experiment demonstrate that probe-list memory can contribute to performance in the probe-word recognition task, they do not rule out the possibility that sentence memory also contributes to the processes of memory access measured by the response times. A safe conclusion at this point would be that the response times result from a mixture of memory access processes based on probe-list memory and on sentence memory. Determining whether the pattern of results in specific probe-word experiments is due to probe-list memory or to sentence memory requires examining the procedures used in those experiments.

Table 1
Results of Experiment 1

	Type of response			
Probe condition	Positive	% correct	Negative	% correct
Names only	1,040	93.9	1,103	95.2
Mixed	1,274	93.9	1,327	94.8

Note. This table shows mean response times (ms) for correct responses to name probes as a function of whether participants were probed only for names or were probed both for names and for other content words.

Experiment 2

The goal of this experiment was to replicate a basic pattern of results that has been found in probe-word recognition studies of co-reference; subsequent studies examined whether this pattern could be attributed to processes other than the establishment and representation of co-reference. The experiment, modeled on one by Gernsbacher (1989), used the same sentences as in Experiment 1. However, unlike Experiment 1, the co-referential expression in the second clause could be either a pronoun or a repeated name as shown in Example 2:

(2) Bill handed John some tickets to a concert but he/Bill took the tickets back immediately.
Antecedent probe: *BILL*
Nonantecedent probe: *JOHN*

Following Gernsbacher again, we manipulated the target of the probe word on positive trials: Antecedent probes (*BILL* in the example) were the names referred to in the second clause of the sentence by either the pronoun or by the repeated name, whereas nonantecedent probes (*JOHN* in the example) were the other names. Gernsbacher found that repeated-name co-reference caused faster response times for antecedent probes than did pronominal co-reference. She interpreted this facilitation of response times as occurring because the repeated name enhanced the matching memory representation created by the first occurrence of that name. She found the opposite pattern for nonantecedent probes: Response times were elevated by repeated-name co-reference as compared with pronominal co-reference. She interpreted this inhibition as occurring when the repeated name suppressed the memory representation of the other, nonmatching name. Experiment 2 sought to replicate that basic pattern of facilitation and inhibition of response times by repeated names that Gernsbacher observed.

Method

Participants. Forty participants from the same population as in Experiment 1 served in the experiment.

Stimulus materials, design, and procedure. The sentences from Experiment 1 were modified for this experiment. As shown in Example 2, four conditions were created by the combination of two factors: form of co-referential expression (repeated name vs. pronoun) and target of the probe word (antecedent vs. nonantecedent). Following Gernsbacher (1989), we used proper names for all of the probe words on both positive and negative trials. Otherwise, the design and procedure were the same as in Experiment 1.

Results and Discussion

The results are shown in Table 2, which gives the response times for accurate responses to antecedent and nonantecedent probes as a function of the form of the co-referential expression (name vs. pronoun). The effect of the form of the co-referential expression in the second clause was not significant, $F_1 < 1$, $MSE = 20{,}612$, $F_1 < 1$, $MSE = 16{,}741$.

Table 2
Results of Experiment 2

	Form of co-referential expression			
Probe type	Repeated name	% correct	Pronoun	% correct
Antecedent	946	98.4	1,025	93.6
Nonantecedent	1,126	90.5	1,049	94.5

Note. This table shows mean response times (ms) for correct responses to probes as a function of the form of the co-referential expression in the second clause and whether the probe matched the antecedent or nonantecedent of the co-referential expression.

Response times were significantly faster for antecedent probes than for nonantecedent probes, $F_1(1, 39) = 55.3$, $MSE = 15{,}242$, $p < .001$, $F_2(1, 55) = 35.4$, $MSE = 15{,}540$, $p < .001$. However, this effect was modified by a significant interaction between the form of the co-referential expression in the second clause and whether it matched the antecedent or the nonantecedent probe, $F_1(1, 39) = 20.9$, $MSE = 23{,}555$, $p < .001$, $F_2(1, 55) = 32.5$, $MSE = 10{,}114$, $p < .001$. The nature of the interaction was the same as that observed previously by Gernsbacher (1989). Response times to antecedent probes were faster in the name condition than in the pronoun condition (facilitation), whereas response times to nonantecedent probes were slower in the name condition than in the pronoun condition (inhibition). The pattern of accuracy to probes mirrored that of response times. The form of the co-referential expression in the second clause did not affect accuracy significantly. Accuracy was higher for antecedent probes than for nonantecedent probes, $F_1(1, 39) = 12.0$, $p < .001$, $F_2(1, 55) = 9.8$, $p < .005$, and there was a significant interaction between the form of the co-referential expression in the second clause and whether it matched the antecedent or the nonantecedent probe, $F_1(1, 39) = 17.8$, $p < .001$, $F_2(1, 55) = 23.3$, $p < .001$.

The pattern of facilitation and inhibition observed in this experiment replicates the pattern found by Gernsbacher (1989) for probes presented at the end of sentences that contained pronouns whose gender did not unambiguously identify a referent, which were the conditions studied in the current experiment. Gernsbacher also found similar patterns for probe words presented immediately after the co-referential expression when pronoun gender was either ambiguous or unambiguous. For present purposes, our replication of Gernsbacher's results shows that our methods are sufficiently similar to hers as to provide a basis for analyzing the phenomena that she uncovered concerning the effects of repeated names on responses to probe words.

Experiments 3a and 3b

These experiments examined whether the establishment of co-reference is a necessary condition for the pattern of inhibition and facilitation in response times to the probe word that is produced by name repetition. We did so by pairing repetitions of a first name with two different last names as shown in Example 3. By the conventions of

English naming, the different last names indicate that the two occurrences of the first name (i.e., *BILL* in the example) refer to different individuals and therefore are not co-referential. The sentence-memory assumption and the probe-list hypothesis make different predictions about the results of the experiment. Following the sentence-memory assumption, the pattern of facilitation and inhibition in probe response times is caused by the mechanisms for establishing co-reference, as Gernsbacher (1989) argued. Accordingly, the addition of different last names should eliminate the pattern because the repeated names are no longer co-referential. In contrast, the probe-list hypothesis asserts that the pattern of facilitation and inhibition is due to two occurrences of a first name on the list of possible probes of which participants are keeping track. Because this list has nothing to do with co-reference, the addition of different last names should not affect the pattern of response times.

Two versions of the experiment were conducted. Experiment 3a compared sentences with repeated first names but different last names, like that in Example 3, to sentences with pronouns, like that in Example 4. This comparison allowed examination of whether the difference between repeated names and pronouns observed by Gernsbacher (1989) and replicated in Experiment 2 occurs when the use of different last names indicates that the second occurrence of the name is not co-referential with the first. Experiment 3b compared sentences with repeated first names but different last names, like that in Example 3, to sentences with repeated first and last names, like that in Example 5. This comparison allowed direct examination of whether co-referential repeated names and noncoreferential repeated names have different effects on probe-word response times.

(3) Bill Whitaker handed John Childress some tickets to a concert but Bill Darden said the tickets were counterfeit.
Repeated name probe: *BILL*
Nonrepeated name probe: *JOHN*

(4) Bill Whitaker handed John Childress some tickets to a concert but he took the tickets back immediately.
Antecedent probe: *BILL*
Nonrepeated name probe: *JOHN*

(5) Bill Whitaker handed John Childress some tickets to a concert but Bill Whitaker took the tickets back immediately.
Repeated name probe: *BILL*
Nonrepeated name probe: *JOHN*

Method

Participants. Eighty participants (40 in each version of the experiment) served in the experiments. They were from the same population as in Experiment 1.

Stimulus materials, design, and procedure. The sentences from Experiment 2 were adapted for this experiment. For the pronoun and the repeated-first-and-last-name conditions, the sentences were the same as in Experiment 2 except that last names always accompanied first names. Last names were selected from the Chapel Hill phone book with the constraints that names varied across the alphabet in initial letter, had multiple entries (and were therefore likely familiar), and were not obviously idiosyncratic to the region in the judgment of the experimenters. For the repeated-first-name-different-last-name condition, the second clauses of the sentences were rewritten so that they made sense with the new characters and were not ambiguous. This rewriting was done following the model, repeated here as Example 3, provided in Experiment 6 of Gernsbacher (1989), which explored the effect on probe response times of introducing a new name in the second clause of the sentences. Each experiment included four conditions created by combining the two types of sentences with probe words that matched the nonrepeated first name or the repeated first name. Correct answers to the postprobe comprehension question required knowledge of both first and last names. Other aspects of the stimuli, design, and procedure were the same as in the preceding experiment.

Results and Discussion

The mean response times for accurate responses in Experiment 3a are shown in Table 3; the pattern of these response times is very similar to that observed in Experiment 2, which involved co-referential names, even though in Experiment 3a, the repeated first names were not co-referential because of different last names. The effect of the type of referential expression in the second clause was marginally significant, $F_1(1, 39) = 4.4$, $MSE = 28{,}602$, $p < .05$, $F_2(1, 55) < 1$, $MSE = 24{,}411$. Response times were significantly faster for probes in the repeated-name/antecedent condition than in the nonrepeated-name condition, $F_1(1, 39) = 30.5$, $MSE = 31{,}212$, $p < .001$, $F_2(1, 55) = 30.7$, $MSE = 21{,}953$, $p < .001$. However, this factor entered into a significant interaction with the type of referential expression, $F_1(1, 39) = 28.6$, $MSE = 31{,}289$, $p < .001$, $F_2(1, 55) = 6.8$, $MSE = 62{,}895$, $p < .02$. This is the same pattern of reaction times that was observed in Experiment 2, showing that nonco-referential repeated names produce a very similar pattern of response times to co-referential repeated names. Accuracy in the probe task was greater in the repeated-name/antecedent condition than in the nonrepeated-name condition, $F_1(1, 39) = 77.3$, $p < .001$, $F_2(1, 55) = 54.4$, $p < .001$. However, this factor entered into a significant interaction with the type of referential expression, $F_1(1, 39) = 38.9$, $p < .001$, $F_2(1, 55) = 34.1$, $p < .001$.

The mean response times for accurate responses in Experiment 3b are shown in Table 3; the pattern of these times indicates that the critical factor is whether a first name is repeated, not whether that name is co-referential as determined by the last name. The effect of the type of referential expression in the second clause was not significant, $F_1(1, 39) < 1$, $MSE = 27{,}211$, $F_2(1, 55) < 1$, $MSE = 41{,}464$. Once again, response times were significantly faster for probes that matched the repeated first names as compared with probes that matched the nonrepeated first names, $F_1(1, 39) = 52.6$, $MSE = 75{,}547$, $p < .001$, $F_2(1, 55) = 178.6$, $MSE = 15{,}132$, $p < .001$. There was not a significant interaction between this factor and whether the last name that accompanied the repeated first name was the same or different.

Taken together, the results of Experiments 3a and 3b indicate that the establishment of co-reference is not a

Table 3
Results of Experiments 3a and 3b

Probe type	Repeated first name–different last name	% correct	Pronoun (Experiment 3a) or repeated first name–same last name (Experiment 3b)	% correct
Type of referential expression: Experiment 3a				
Repeated name–antecedent	1,066	96.7	1,135	88.4
Nonrepeated name	1,283	77.4	1,147	86.1
Form of co-referential expression: Experiment 3b				
Repeated name	1,033	96.7	1,023	94.6
Nonrepeated name	1,247	74.6	1,257	71.0

Note. This table shows mean response times (ms) for correct responses to probes as a function of the type of referential expression in the second clause and whether the probe matched the name that was repeated (or was the antecedent) or the name that was not repeated.

necessary condition for the inhibitory and facilitative effects of repeated names on probe-word response times observed by Gernsbacher (1989) and replicated here in Experiment 2. They do so by showing that pairing two occurrences of a first name with different last names does not change the pattern of response times from that observed when no last names are used or when the same last name is used with both occurrences of a first name.

Experiment 4

The preceding experiments have established that the probe-word recognition task is influenced by strategies for keeping track of words that are likely to be probed (Experiment 1) and that the effects of name repetition on response times to probe words occur even when the use of different last names makes it unlikely that name repetition results in co-reference (Experiments 3a and 3b). This raises the possibility that the effects of name repetition do not stem from processing the structure or meaning of a sentence but instead emerge from strategies that treat the names as members of a list to be remembered; that is, the effects emerge from probe-list memory. That notion was tested in Experiment 4 by examining the effects of name repetition when the task was modified so that it did not involve sentence comprehension. This was done by keeping the probe-word task the same but by randomizing the positions of the words in the sentences (other than the key referential expressions), as shown in Example 6.

(6) Bill but John some to tickets a back concert he/Bill handed tickets immediately the took.
Repeated name or antecedent probe: *BILL*
Nonrepeated name probe: *JOHN*

The secondary task was changed from a comprehension test of the meaning of the sentence to determining whether the list contained a misspelled word. This was done in order to make sure that participants paid attention to all the words in the sequence. Thus, if the effect of name repetition on probe-word response times can be caused by determining which words on a list are likely to be probed and does not require comprehension of the structure or meaning of a sentence, then the pattern observed by Gernsbacher (1989) and in Experiments 2 and 3a should be seen in the Experiment 4 even though it does not require comprehension of a sentence.

Method

Participants. Forty participants from the same population as in the previous experiments served in this one.

Stimulus materials, design, and procedure. Stimuli were constructed from the sentences used in Experiment 2. The positions of the words, other than the names and the co-referential expression (repeated name or pronoun), were randomized. In half of the lists, a letter was changed in one of the words to create a nonword. The position of the nonword was randomly varied over the last four positions in the word list. This ensured that participants had to attend to the list until all names and/or pronouns had occurred. Participants were instructed in the probe-word task as they had been in the previous experiments. After they had responded to the probe word, they were required to make a nonspeeded response to a question about whether there was a misspelled word in the list. Other aspects of the design and procedure were the same as in the previous experiments.

Results and Discussion

Table 4 shows the mean response times (and accuracy rates) for correct responses in the probe-word task. The main effect of the type of the third referential expression (name vs. pronoun) was not significant, $F_1(1, 39) < 1$, $MSE = 15{,}141$, $F_2(1, 55) < 1$, $MSE = 11{,}885$. Response times were significantly faster for probes in the repeated-name/antecedent condition than in the nonrepeated-name condition, $F_1(1, 39) = 22.1$, $MSE = 14{,}248$, $p < .001$, $F_2(1, 55) = 17.9$, $MSE = 12{,}337$, $p < .001$. However, this factor interacted significantly with whether the third referential expression was a repeated name or a pronoun, $F_1(1, 39) = 28.9$, $MSE = 11{,}127$, $p < .001$, $F_2(1, 55) = 14.2$, $MSE =$

Table 4
Results of Experiment 4

	Form of second expression			
Probe type	Repeated name	% correct	Pronoun	% correct
Repeated name–antecedent	994	97.7	1,051	93.6
Nonrepeated name	1,121	91.4	1,050	93.8

Note. This table shows mean response times (ms) for correct responses to probes as a function of the form of the second expression and the type of probe. In the repeated-name condition, probes could match either the repeated name or the nonrepeated name. In the pronoun condition, the probe could match either the name that was the antecedent of the pronoun before the sentence was randomized or the name that was not the antecedent.

18,305, $p < .001$. The pattern of accuracy in the probe task mirrored the pattern of response times. There was no significant difference in accuracy as a function of the type of the third referential expression (name vs. pronoun). Accuracy was significantly higher in the repeated-name/antecedent condition than in the nonrepeated-name condition, $F_1(1, 39) = 10.3, p < .01, F_2(1, 55) = 8.9, p < .005$. This factor interacted significantly with whether the third referential expression was a repeated name or a pronoun, $F_1(1, 39) = 8.9, p < .005, F_2(1, 55) = 6.3, p < .025$.

Compared with a pronoun, a repeated name in the stimulus list caused faster response times to probes for that name and slower response times for the nonrepeated name. This is the same pattern that was observed in Experiment 2 in which participants engaged in a language comprehension task involving co-reference between names, and it was observed in Experiment 3a in which participants engaged in a language comprehension task involving disjoint reference between two names. Observation of this pattern in a task that involves processing a random sequence of words rather than comprehending a sentence shows that the effect can be caused by processes of memory representation and retrieval in a list of words and does not require accessing the memory representation of the meaning of a sentence.

General Discussion

The experiments in this article have addressed the questions of whether response times in probe-word recognition tasks are influenced by strategies aimed at keeping track of the words that are likely to be probed and of whether such strategies may be sufficient to account for the effects of repeated names on response times to probe words. These are effects that have been previously attributed to the mechanisms for establishing co-reference. Experiment 1 showed that response times to probe words were influenced strongly by the number of words that could possibly serve as probes. This finding is consistent with the hypothesis that participants formed a memory representation of the list of words that were likely to be probed. Increases in the length of the probe list led to increases in response times, a common finding in memory research on lists. The subsequent experiments examined whether the effects of repeated names on response times, effects that have been previously interpreted as a reflection of the mechanisms of referential processing, can be attributed to probe-list memory. Experiment 2 replicated a basic pattern of facilitation and inhibition of response times by name repetition in sentences with co-reference. Experiments 3a and 3b showed that these effects occurred whether or not the repeated name was co-referential with a prior name (because of the presence of a same or different last name), and Experiment 4 showed that the effects occurred when the order of the words in a sentence was randomized and the task did not involve language comprehension at all. These results indicate that the effects of name repetition can be attributed to probe-list memory and therefore that the effects should not be interpreted as providing evidence about how reference is established. This conclusion weakens models (e.g., Gernsbacher, 1989, 1990, 1996) in which repeated names are characterized as more effective than pronouns in establishing co-reference and strengthens models (e.g., Gordon & Hendrick, 1998) that present the opposite view about the relative effectiveness of the two types of forms in establishing co-reference.

Whereas the current experiments indicate that probe-list memory may be sufficient to account for the effects of name repetition, they do not indicate why repeating a name on a list causes faster response times to the repeated name and slower response times to a nonrepeated name. Research on recognition memory offers some suggestions. In particular, Murnane and Shiffrin (1991) interpreted effects on recognition-memory performance of repeating words in different sentences as indicating that separate memory representations were created for the repetition, resulting in an increase in the length of the memory list. The existence of two memory representations for the repeated word enhances its subsequent recognition, but the increase in list length impairs recognition of words that were presented only once (Murnane & Shiffrin, 1991). Although it is unclear whether the model advanced by Murnane and Shiffrin is a valid account of the effects of name repetition in probe-word recognition tasks, the parallel between the two sets of findings suggests that the phenomena addressed in list-learning experiments and in probe-word recognition studies of reference may be similar.

In the present article, evidence that probe-list memory is sufficient to account for the effects of name repetition was obtained in experiments that probed only for names. Researchers using the probe-word task to study referential processing have typically used probes consisting of proper names in the conditions of experimental interest. Some researchers have exclusively used name probes both as lures and in filler items (e.g., Gernsbacher, 1989), whereas other researchers have included some nonname probes in filler items (e.g., Chang, 1980; Greene et al., 1992). It is an open question whether including nonname probes as filler items can forestall task-specific strategies. The answer probably

Table 5
Proportions of Probe Types for Positive and Negative Trials

Variable	Positive Name (%)	Positive Content (%)	Negative Name (%)	Negative Content (%)	Proportion of sentence
Chang (1980)					
Experiment 2	65	35	41	59	.20[a]
Corbett & Chang (1983)					
Experiment 1	67	33	46	54	.22[b]
Experiments 2–3	59	41	46	54	.27[b]
Gernsbacher & Hargreaves (1988)					
Experiments 1–7	100	0	100	0	.21–.28[a]
Gernsbacher (1989)					
Experiments 1–6	100	0	100	0	.18[b]
Gernsbacher, Hargreaves, & Beeman (1989)					
Experiments 1–6	100	0	100	0	.20–.29[a]
MacDonald & MacWhinney (1990)					
Experiments 1–2	67	33	?	?	.10[c]
Greene, McKoon, & Ratcliff (1992)					
Experiments 1–4, 7	63	37	40	60	.12[c]
Experiment 5	100	0	100	0	.11[c]
Experiment 6	67	33	74	26	.11[c]
Experiments 8–9	50	50	?	?	.11[c]
McKoon, Greene, & Ratcliff (1993)					
Experiments 1–6	44	56	49	51	.13[c]
Experiment 7	56	44	49	51	.12[c]
McDonald & MacWhinney (1995)					
Experiments 1–3	67	33	67	33	.20[b]
Carreiras, Gernsbacher, & Villa (1995)					
Experiments 1–2	100	0	100 ?	0	.25[a]
Experiment 3	50	50	50	50	.25[a]
Carreiras (1997)					
Experiment 3	100	0	100	0	.08[a]
Garnham, Traxler, Oakhill, & Gernsbacher (1996)					
Experiments 1–4	100	0	100	0	.18[b]

Note. The Name and Content columns show the proportion of probes that are names and the proportion of probes that are other content words for positive trials (probe present in language stimulus) and for negative trials (probe absent in language stimulus) from articles that used names as probes in experimental conditions. Question marks indicate instances in which either we were not confident of our interpretation of the method section, or the information was omitted. The final column gives the proportion of the words in the language stimulus that were proper names.
[a]Proportion (or range of proportions across experiments) estimated from sample(s) in text. [b]Proportion estimated from five stimuli in the appendix of the cited article. [c]Proportion based on information provided in text.

depends on whether the kinds of words being probed for are predictable by readers. Table 5 shows the proportion of names and other content words used as probes in studies of the referential processing of names, as well as estimates of the proportion of content words in the sentences (or other language fragment) that were names. In all cases, the proportion of probes that were names exceeds the proportion of the words in sentences that were names. This suggests that using probe-list memory to keep track of names in these studies might be advantageous for performance on the probe-word recognition task.

In addition to being used to study the establishment of co-reference with repeated names, the probe-word recognition task has also been used to study the processing of pronouns. This raises the question of whether effects of pronouns on probe-word recognition can be accounted for by probe-list memory. Because pronouns achieve reference only through their relation to other expressions in a sentence or passage, it is not clear how processing exclusively associated with probe-list memory could result in a pronoun influencing response times to a probe-word name. Perhaps it could be argued that when a pronoun is interpreted as a result of language comprehension, the strategies associated with the probe-word task induce the participant to update the probe-list representation of the referred-to name. By this argument, probe-word responses would again be based on probe-list memory rather than on sentence memory. This argument seems difficult to refute, but it also does not seem critical because it concedes that effects of pronouns on probe-word response times are a marker of pronoun interpretation. The more critical point is that pronouns tend not to cause effects on probe-word response times. Both of our experiments that included co-referential pronouns (Experiments 2 and 3a) found that the probe-word response times did not differ significantly for antecedent and nonantecedent probes. Table 6 lists the experiments that have investigated

Table 6
Published Experiments Examining the Effects of Co-Referential Pronouns on Response Times to Probe Words Consisting of Proper Names

Variable	Published experiments
Pronoun affects probe response time	Carreiras (1997), Experiment 3 Chang (1980), Experiment 2 Corbett & Chang (1983), Experiment 1 Garnham, Traxler, Oakhill, & Gernsbacher (1996), Experiment 4[a] Gernsbacher (1989), Experiments 4 & 5 (tested at end of sentence) Greene, McKoon, & Ratcliff (1992), Experiments 5 & 6 MacDonald & MacWhinney (1990), Experiments 1 & 2 McKoon, Greene, & Ratcliff (1993), Experiments 1–4
Pronoun does not affect probe response time	Corbett & Chang (1983), Experiments 2 & 3 Garnham, Traxler, Oakhill, & Gernsbacher (1996), Experiments 1–3[a] Gernsbacher (1989), Experiments 1–3 Gernsbacher (1989), Experiments 4 & 5 (tested immediately after pronoun) Greene, McKoon, & Ratcliff (1992), Experiments 1–4, 7–9 (relative to control) McDonald & MacWhinney (1995), Experiments 1 & 3[a] McKoon, Greene, & Ratcliff (1993), Experiments 5–7

Note. Experiments in this table are grouped according to whether they provided evidence that the pronoun facilitated response times to antecedent probe names as compared with nonantecedent probe names.
[a]This particular set of experiments (Garnham, Traxler, Oakhill, & Gernsbacher, 1996; McDonald & MacWhinney, 1995) focused on the effects of the implicit causality of verbs on the activation of referents, and the pronoun effects (or absence of effects) should be seen as relating to the impact of verb causality.

the effects of pronouns on responses to antecedent and nonantecedent probes; the majority of experiments have found no effect. Greene et al. (1992) have argued that the absence of pronoun effects on probe-word response times indicates that pronouns are not automatically interpreted as part of normal language comprehension. An alternative interpretation is that performance in probe-word tasks is usually determined by probe-list memory on which normal language comprehension processes have little effect. This might explain why the effects of pronouns on probe-word responses are inconsistent and vary with subtle differences in the experimental task.

A final domain in which to consider the implications of probe-list memory concerns studies that have used content words that are not names as probes in the conditions of experimental interest. Again, we argue that the likelihood that responses are based on probe-list memory rather than on sentence memory would be strongly influenced by whether the words that are likely to be probed are predictable by readers. There is considerable variety in the types of language stimuli that have been used in experiments with content words as probes (e.g., McKoon & Ratcliff, 1992; McKoon, Ward, Ratcliff, & Sproat, 1993), so it is difficult to offer a broad comparison of those studies like that presented for the predictability of names in Table 5. However, it is worth noting that some of the studies using content words as probes have used relatively lengthy language stimuli (e.g., 61 words in the example shown in McKoon & Ratcliff, 1992, and 62 words in the example shown in McKoon et al., 1993). The length of the stimuli could either diminish or enhance the attractiveness of pursuing a probe-list strategy in performing the task. On the one hand, keeping track of all the words in a separate probe-list memory might overwhelm the capacity to keep track of unrelated words, thereby making reliance on sentence memory more likely. On the other hand, memory for literal expressions in long passages is not good, and there could be a real advantage in pursuing a probe-list strategy if the structure of the task implicitly provides information about which words are likely to be probed.

The probe-word recognition task has played a substantial role in research on language comprehension. Some of the researchers (e.g., Gernsbacher, 1989; Greene et al., 1992; McKoon, Gerrig, & Greene, 1996) who have used probe-word methodology to study referential processing have argued that their results show that language comprehension can be explained by general models of information representation and retrieval that have been developed to explain data obtained in memory and word recognition experiments (Gillund & Shiffrin, 1984; McClelland & Rumelhart, 1981; Murdock, 1982; Ratcliff, 1978). The present results suggest that at least part of the ease of explaining probe-word results using general models of memory may derive from the fact that in some cases the probe-word recognition task induces participants to create the kinds of representations that are also useful in memory experiments testing lists of words. Experiments that use probe-word recognition to study language comprehension should include controls that support the conclusion that the specific results derive from processes and representations involved in language comprehension and not from processes and representations that derive from strategic adaptations to the probe-word recognition task.

References

Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception & Psychophysics, 12,* 73–76.

Carreiras, M. (1997). Plural pronouns and the representation of their antecedents. *European Journal of Cognitive Psychology, 9,* 53–87.

Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review,* 2(1), 124–129.

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

Cloitre, M., & Bever, T. G. (1988). 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., Ratcliff, R., & McKoon, G. (1981). Study and test repetition effects in item recognition priming. *American Journal of Psychology, 94,* 497–511.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory & Language, 35,* 517–543.

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 & 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.

Gernsbacher, M. A. (1996). The Structure-Building Framework: What it is, what it might also be, and why. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 289–312). Mahwah, NJ: Erlbaum.

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

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

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 19,* 1–65.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17,* 311–347.

Gordon, P. C., & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition, 62,* 325–370.

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science, 22,* 389–424.

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 266–283.

Hockley, W. E., & Corballis, M. C. (1982). Tests of serial scanning in item recognition. *Canadian Journal of Psychology, 36,* 189–212.

Hockley, W. E., & Murdock, B. B. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review, 94,* 341–358.

Hudson-D'Zmura, S. B. (1988). *The structure of discourse and anaphore resolution: The discourse center and the roles of nouns and pronouns.* Unpublished doctoral dissertation, University of Rochester.

Hudson, S. B., Tanenhaus, M. K., & Dell, G. S. (1986). The effect of the discourse center on the local coherence of a discourse. In *Proceedings of the 8th Annual Conference of the Cognitive Science Society* (pp. 96–101). Hillsdale, NJ: Erlbaum.

Kennison, S. M., & Gordon, P. C. (1997). Comprehending referential expressions during reading: Evidence from eye tracking. *Discourse Processes, 24,* 229–252.

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

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

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

McKoon, G., Gerrig, R., & Greene, S. (1996). Pronoun resolution without pronouns: Some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 919–932.

McKoon, G., Greene, S. B., & Ratcliff, R. (1993). Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1040–1052.

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

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440–466.

McKoon, G., Ward, G., Ratcliff, R., & Sproat, R. (1993). Morphosyntactic and pragmatic factors affecting the accessibility of discourse entities. *Journal of Memory & Language, 32,* 56–75.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89,* 609–626.

Murnane, K., & Shiffrin, R. (1991). Word repetitions in sentence recognition. *Memory & Cognition, 19,* 119–130.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R., & McKoon, G. (1981). Automatic and strategic priming in recognition. *Journal of Verbal Learning & Verbal Behavior, 20,* 204–215.

Sternberg, S. (1966, August). High-speed scanning in human memory. *Science, 153,* 652–654.

Received April 12, 1999
Revision received August 16, 1999
Accepted August 25, 1999 ■