Morphosyntactic and Pragmatic Factors Affecting the Accessibility of Discourse Entities

GAIL McKOON, GREGORY WARD, AND ROGER RATCLIFF

Northwestern University

AND

RICHARD SPROAT

AT&T Bell Laboratories

Six experiments provide results showing that the accessibility of discourse entities is affected jointly by pragmatic and morphosyntactic factors. Accessibility was varied pragmatically by making an entity more or less closely related to the topic of its discourse, and it was varied syntactically by introducing an entity either in a verb phrase (*deer* in *hunting deer*) or in a compound (*deer hunting*); the latter should be less accessible according to linguistic data. The accessibility of an entity was examined by measuring the difficulty of understanding a pronoun intended to refer to the entity. Difficulty of understanding the pronoun was measured with reading time for a sentence mentioning the entity, with a test of short term memory, and with a test of long term memory. Results showed that both the pragmatic and syntactic variables affected reading time for the sentence with the pronoun, but that in all cases the relationships among the referent, the pronoun, and information given in the discourse about them appeared to be understood both in their representation in short term memory and in their representation in long term memory. © 1993 Academic Press, Inc.

An important aspect of understanding language, whether listening to a speaker or reading a text, is relating each new piece of information to information that has already been conveyed. This context of prior information is assumed to be represented in "working memory" and used in determining the meanings of individual words, the relations among individual propositions, and the relevance of concepts and propositions to the overall message. The information in working memory is especially critical for the interpretation of pronouns and other anaphoric expressions. In this article,

we investigate the structure of information in working memory as it relates to the comprehension of pronouns. We assume a complex structure that is determined by both morphosyntactic and pragmatic factors; following recent work in computational linguistics and discourse analysis, we label this structure a "discourse model." In six experiments, we investigate some of the referential properties of such a model. The experiments investigate the ease with which specific entities in the discourse model may be accessed by means of pronominal reference, and they show that successful reference is a function of both the pragmatic and syntactic context in which the referent was evoked in the prior discourse.

Within cognitive psychology, there have been two distinct traditions of text processing research that have investigated how on-line language comprehension in general, and anaphor interpretation in particular, re-

56

late to the representation of information in a discourse model. One tradition has generally focused on syntactic determinants of linguistic structure, and, more narrowly, on structure within a single sentence. Under this view, the relationships among the elements of a sentence are organized according to the syntactic roles that they fill in that sentence. Reference to concepts or entities previously evoked by the text is accomplished by accessing syntactically defined elements; an anaphor accesses the syntactic part of the sentence in which its antecedent occurs. Ease of access is determined by the position of the antecedent in the syntactic structure. Mathews and Chodorow (1988), for example, provide data suggesting that antecedents more deeply embedded in a syntactic structure lead to more difficulty for the interpretation of an anaphor than antecedents not so deeply embedded. In a similar vein, data from experiments by Nicol and Swinney (1989) suggest that the availability of a potential referent is a function of its "syntactic appropriateness" as the antecedent of an anaphor. Syntactic approaches to the on-line representation of discourse information are reviewed by Mathews and Chodorow (1988) and by Fodor (1989).

The other traditional approach to the on-line processes and representations relevant to anaphora has focused on the structure of a discourse as a whole, rather than on single sentences (cf. Haviland & Clark, 1974; Malt, 1985). Kintsch (1974) proposed that a discourse was made up of semantic propositions ("individual idea units") and that these propositions were connected to each other through shared arguments. A connected set of propositions was assumed to consist of a "topic proposition," i.e., the most important proposition of the set, and the importance of all other propositions was defined relative to this proposition. Kintsch and van Dijk (1978) later incorporated this structural proposal into a model of on-line comprehension. In this model,

each new set of propositions in a discourse is added to the already existing structure via connections among shared arguments, with preference given to more recently mentioned propositions and arguments. Entities of the discourse that are more topical are more likely to be kept active in short-term memory, and therefore they are more available as referents of anaphoric elements.

The "discourse model" approach that we assume as the background for our research combines elements from the two traditions in psycholinguistics and from computational linguistics, and also introduces several new elements. Following Sidner (1981), Webber (1979), and the propositional tradition (Haviland & Clark, 1974; Kintsch, 1974), we assume that discourse models contain the entities ("arguments," Kintsch, 1974, or "cognitive elements," Sidner, 1981) evoked in a discourse, and these entities are linked together by the relations in which they participate. The entities in question are assumed to be *conceptual* entities—not linguistic ones. As Morgan (1978), Webber (1979), Sidner (1981), and others have pointed out, language and, in particular, referring expressions, are used to refer to objects in the world (or model thereof), and not to other linguistic units.

We also assume that the entities represented in the discourse model are associated with varying degrees of accessibility. Not all noun phrases evoke discourse entities. For example, the anaphor *it* in the sentence *It's snowing outside* does not evoke a discourse entity (cf. Kamp, 1981; Heim, 1982; Webber, 1983), and so the notion of accessibility does not apply. Other anaphors, such as *do so*, have been argued to require explicit linguistic antecedents (McKoon et al., in preparation; Murphy, 1985; Tanenhaus & Carlson, 1990) and therefore may be more sensitive to surface form than to the discourse level of representation. In this article we exclude these

kinds of anaphors and restrict discussion to anaphors that are used to evoke discourse entities in a discourse model and consider their varying degrees of accessibility. We assume that the entire current discourse—and not just individual component sentences—is represented in the discourse model (cf. Kintsch, 1988), although at times, of course, portions of it will be relatively inaccessible and other portions will be particularly salient, or "in focus" (cf. Grosz, 1978; Grosz & Sidner, 1986). Which entities are highly accessible ("in focus") will change as the discourse progresses, partly as a function of recency, and partly as a function of shifts in topic (cf. Malt, 1985).

Our notion of a discourse model differs from previous psycholinguistic proposals in two key ways. First, we claim that the accessibility of discourse entities for subsequent anaphoric reference is determined not by syntax alone and not by topicality alone, but by a variety of syntactic, pragmatic, and semantic factors. The critical consequence of this claim is that there need be no single, most accessible entity (such as the topic) in the discourse, nor is there a single metric (such as syntactic depth of embedding) by which accessibility can be calibrated. Experiments 1 through 6 support this claim by showing that accessibility depends simultaneously on both syntactic and pragmatic factors.

Second, we maintain that the accessibility of an entity in a discourse model is determined not only by the context in which it is introduced but also by the cue with which that entity is later accessed by the comprehension system. Different cues may access the same entity with varying degrees of success; in some contexts, a definite description may work better than a pronoun, and in other contexts, the reverse might be true. Furthermore, the entities that are most accessible given one cue may be different from the entities that are most accessible given another cue. For example, a pronoun may serve to evoke more recent entities, whereas a definite description might serve to evoke more distant entities. Our notion is that reference processing is an interaction between an anaphoric cue and discourse entities in memory. Later in this article, we describe this notion through the metaphor of current global memory models and show how it guides the methodology used in the experiments.

It is important to note the limitations on the theoretical discourse model that we assume. The model is hypothesized to include entities that are explicitly mentioned in the discourse, the relations among those entities (cf. Kintsch, 1974), and their accessibilities relative to potential cues. Whether information of other kinds, such as inferences, "mental models," or causal structures, is also included in the working memory representation of text is an open question (McKoon & Ratcliff, 1992). Thus, for present purposes, our conception of a discourse model represents only the information necessary for processing the kinds of anaphora under investigation, and therefore it differs from the models that have been proposed by some other researchers (Bransford, Barclay, & Franks, 1972; Johnson-Laird, 1983; Morrow, Bower, & Greenspan, 1989; Oakhill, Garnham, & Vonk, 1989; Sanford & Garrod, 1981).

Because the discourse model theory assumed in our research contains elements of previous approaches, it is consistent with a number of previous empirical findings. In Kintsch's model for on-line text comprehension (Kintsch, 1988), the accessibility of an entity depends on the recency with which it was evoked and on how closely connected it is to the discourse topic. Empirically, both of these variables have been demonstrated to affect accessibility as hypothesized: it has been shown that more recently mentioned entities are more accessible (Jarvella, 1971; Caplan, 1972), and that entities more closely connected to the topic are better recalled (Kintsch &

Keenan, 1973) and better recognized (McKoon, 1977). Because the discourse model theory incorporates both recency and topicality as variables affecting accessibility, these findings are consistent with it.

The theory is also consistent with research motivated by more syntactic views of discourse representation. Under these views, the accessibility of an anaphor for an antecedent depends on the syntactic position of the antecedent. Mathews and Chodorow (1988), for example, tested comprehension of the pronoun in sentences like (1a) and (1b):

(1a). After the bartender served the patron, he got a big tip.
(1b). After the bartender served the patron, he left a big tip.

They found that reading time for the clause with the pronoun was faster when the antecedent of the pronoun occurred in subject position than when it occurred in object position. On a strictly syntactic account, this advantage would be due to a search process for the antecedent through the sentence's syntactic structure. An antecedent in subject position, as in (1a), would have an advantage in a left-to-right or top-down search. A discourse model approach would also predict an advantage when the antecedent is in subject position, but not because of a search through a syntactic structure. Instead, the advantage would be due to the greater accessibility in the discourse model of entities evoked in subject position relative to entities evoked in object position.

In our view of discourse models, syntax is assumed to be one of the factors that determines the relative accessibilities of the entities in the model. Several studies have investigated such effects. Rothkopf, Biesenbach, and Billington (1986) and Rothkopf, Koether, and Billington (1988) have shown that a modifier is better recalled when it is presented in predicate adjective position than when it is presented in prenominal position. In Rothkopf et al.'s experiments, texts contained sentences with phrases like *the yellow fruit* or *the fruit that was yellow*. Subjects were better able to answer a later question about the color of the fruit if they had read the second (predicate adjective) version. McKoon, Ward, Ratcliff, and Sproat (in preparation) demonstrated the same point with a different procedure; they showed that a predicate adjective is better recognized than a prenominal one. For example, the adjective *hostile* was presented in either prenominal or predicate position: *The hostile aunt was intolerant* or *The intolerant aunt was hostile*. Later recognition of the word *hostile* was faster and more accurate when it had been read in predicate adjective position. Similarly, concepts presented in direct object position are better recognized than concepts presented in an indirect object position, again demonstrating the effect of syntactic context on later accessibility (McKoon et al., in preparation).

Previous findings such as those just described show either pragmatic influences on accessibility (e.g. Kintsch & Keenan, 1973) or syntactic influences (e.g. Mathews & Chodorow, 1988). What they do not show is that these factors combine in a discourse to jointly affect accessibility for a single discourse entity. This was one of the goals of the experiments presented in this article. Accessibility was examined through its effects on the ease of comprehension of pronouns; the more accessible an entity, the more easily comprehended should be a pronoun being used to refer to that entity.

A second goal of the experiments was to investigate an interesting case of anaphora that has been the topic of much debate in the linguistics literature. This type of anaphora provided us with the means to manipulate accessibility via the syntactic structure by which an entity was introduced into a discourse.

In this type of anaphora, reference is

made to entities evoked by antecedents that appear within morphologically complex words. In the second sentence of (2) below, the pronoun *it* has as its antecedent *Kal Kan. Kal Kan* appears within the complex word *Kal Kan cat,* where we use the notion of *word* as defined in recent studies in morphology (cf., Matthews, 1974; Mohanan, 1986): a word may consist of a combination of a stem plus some affixes, normally written as a single orthographic word in English, or else may be a *compound* of several stems, often written as multiple orthographic words, as is the case with *Kal Kan cat.*

  2. Patty is a definite Kal Kan cat. Every day she waits for it.

A number of linguistic studies have argued that examples like (3b), in which an antecedent occurs within a compound, are ungrammatical, and so have postulated a grammatical prohibition against complex words containing antecedents for anaphoric elements (e.g., Postal, 1969; Lakoff & Ross, 1972; Simpson, 1983; Mohanan, 1986). In particular, Postal (1969) proposed that no anaphor could have as its antecedent a word that was "part of the sense of" another word. Contrasts such as the one exhibited in (3) (Postal, 1969, p. 230) are claimed to be the result of such a grammatical prohibition:

  3a. Hunters of animals tend to like them.
  3b. Animal hunters tend to like them.

According to Postal, *them* can be interpreted as "referring to" animals in (3a), but not in (3b). In (3b), *animal* is morphologically contained within the compound *animal hunters,* which by Postal's constraint constitutes what is called an "anaphoric island," and cannot by grammatical rule provide the antecedent for *them.*

However, Ward, Sproat, and McKoon (1991) have argued against this position, presenting dozens of examples of felicitous naturally occurring tokens from a variety of oral and written sources. The example in

(2) is one of these tokens; others are given in (4) (the specific sources for the examples are given in Ward et al., 1991):

  4a. *Bush* supporters would stay home, figuring *he*'d already won. (*he* = Bush)
  4b. Call if you're a *small business* owner, or interested in starting *one.* (*one* = a small business)
  4c. For a *syntax* slot, I'd rather see someone with more extensive coursework in *it.* (*it* = syntax)
  4d. We went up to *Constable* country; we stayed in the village *he* was born in. (*he* = Constable)
  4e. Millions of *Oprah Winfrey* fans were thoroughly confused last week when, during *her* show, she emotionally denied and denounced a vile rumor about herself. (*her* = Ophrah Winfrey)
  4f. Our neighbors, who are sort of *New York City*-ites, they have jobs *there . . .* (*there* = New York City)
  4g. Do *parental* reactions affect *their* children? (*their* = parents)

Given that examples such as these occur naturally in spoken and written language, it would appear that word-internal elements *can* serve as antecedents for anaphors, contrary to the claims of Postal and others.

In fact, Ward et al. (1991) argue that there is no *grammatical* constraint preventing word-internal elements from serving as antecedents for anaphors. Rather the felicity of such anaphora is a function of the accessibility of the discourse entity evoked by the word internal element to which the anaphor is intended to refer. Consistent with our assumptions about the representation of entities in a discourse model, we claim that both pragmatic and syntactic factors are relevant for the accessibility of the entity. In other words, the factors involved in determining the felicity of anaphora for anaphoric islands are exactly the same as the factors involved in determining the accessibility of discourse entities in general.

According to Ward et al. (1991), the un-

acceptability of anaphora like that in (3b) is due to the inaccessibility of the relevant discourse entity. As mentioned above, modifiers have been shown to be relatively inaccessible (McKoon et al., in preparation; Rothkopf et al., 1986; Rothkopf et al., 1988) and so, assuming that the word-internal element is functioning as a modifier, word-internal elements should not generally be sufficiently accessible to reference by anaphora.

On the other hand, all of the pragmatic, syntactic, and semantic factors that determine accessibility in a discourse model can conspire, singly or jointly, to make word-internal elements sufficiently accessible to permit subsequent anaphora. For example, discourse entities can increase in accessibility through relevance to the listener or reader; *Sheep farmers tend to like them* was judged acceptable by some members of a New Zealand audience. Ward et al. (1991) point out two further ways in which a discourse entity can become more accessible. One way is through contrast with another discourse entity, as in (5), a quote from President Reagan's 1990 farewell speech:

5. Well, action is still needed. If we're to finish the job, Reagan's Regiments will have to become the BUSH Brigades. Soon *he'll* be the chief, and he'll need you every bit as much as I did.

The other way is through topicality. In a television commercial for Saab, the pronoun *it* in sentence (6) can felicitously refer to the Saab model 9000-CD which was evoked by a word internal to the compound *Saab 9000-CD owners*. Similarly, in the first text in Table 1, the topic of the discourse segment is hunting and the discourse entity corresponding to the referent of the pronoun in the last sentence (i.e., *they/deer*) is closely related to the topic; therefore we would hypothesize that it is relatively accessible.

6. We asked *Saab 9000-CD* owners about *its* road-handling . . .

In sum, we have reason to believe not only that the compound construction illustrated in (3b) serves to render an entity rel-

TABLE 1
EXAMPLES OF TEXTS USED IN EXPERIMENT 1

High topicality, compound
Sam likes the outdoor life. Having grown up in rural Kentucky, he knows a lot about nature and is an expert at fishing and shooting. He goes on hunting trips as often as he can. He used to hunt just small game, like rabbit and quail. However, lately he's taken up deer hunting. He thinks that they are really exciting to track.

Low topicality, compound
Sam has many interests in the outdoors. He's an avid skier, and each winter he takes about a month off from work to ski in Colorado. In the summertime, he visits his parents in Montana where he has a chance to do some mountain climbing. Lately, he's taken up deer hunting. He thinks that they are really exciting to track.

High topicality, verbal complement
Sam likes the outdoor life. Having grown up in rural Kentucky, he knows a lot about nature and is an expert at fishing and shooting. He goes on hunting trips as often as he can. He used to hunt just small game, like rabbit and quail. However, lately he's taken up hunting deer. He thinks that they are really exciting to track.

Low topicality, verbal complement
Sam has many interests in the outdoors. He's an avid skier, and each winter he takes about a month off from work to ski in Colorado. In the summertime, he visits his parents in Montana where he has a chance to do some mountain climbing. Lately, he's taken up hunting deer. He thinks that they are really exciting to track.

Note: Referent noun: deer.

atively inaccessible in some discourse contexts but also that an entity evoked in this construction can be made quite accessible in other discourse contexts. The hypothesis of a joint contribution to accessibility of morphosyntactic and pragmatic factors makes a number of predictions amenable to empirical investigation, which we report on below. To anticipate, in Experiment 1, we varied topicality for entities evoked by antecedents contained in the compound and the corresponding verb phrase constructions, as shown in Table 1. Our prediction was that accessibility for the "referent entity" (*deer* in Table 1) would be increased both by the pragmatic and the syntactic variables; the entity would be more accessible when it was more closely related to the topic and when it was introduced in a verb phrase rather than a compound.

How to Measure Accessibility

Given our notion of a discourse model, accessibility is defined as the ease with which a discourse entity, introduced at one point in a discourse, can be referenced at a later point in the discourse by some cue, such as a pronoun. The empirical goal is to measure accessibility by measuring ease of reference, that is, to measure the ease with which pronouns are understood. This requires at least a minimal model of comprehension processes for pronouns.

In Greene, McKoon, and Ratcliff (1992) and Ward, Sproat, and McKoon (1991), we proposed that a pronoun is completely and correctly understood if its intended referent is sufficiently more highly accessible in the discourse model, relative to the pronoun as a cue, than all other discourse entities. Following current global memory models (Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1982; Ratcliff, 1978; see also Gernsbacher, 1989), a pronoun is assumed to be matched against all entities in the discourse model in parallel. The semantic and grammatical features of the pronoun are matched against the features of the discourse entities. Every entity in the discourse model will match the pronoun to some degree, with the degree of match depending on both the entity's semantic and grammatical features and its accessibility. If the degree of match for some single entity is sufficiently high, and sufficiently higher than the match for all other entities, then (without further processing) that entity is identified as the pronoun's referent; in essence, a sufficiently high degree of match constitutes a decision about the pronoun's referent. If there is no entity that matches sufficiently well, then a referent is not identified. If more than one entity matches sufficiently (but none sufficiently better than the others), then again no single referent is identified. In the cases where a referent is not identified, comprehension may fail in the sense that the pronoun is left without a referent. Alternatively, selection of a referent might be postponed, waiting for more information from the discourse, or for strategic problem solving processes that might be able to identify a referent. In the usual case, where a single entity matches the pronoun sufficiently better than all other entities, the identification of the pronoun with the referent leads to the attachment in the discourse model of information associated with the pronoun to information associated with the referent.

This model for comprehension of pronouns makes the explicit claim that pronouns vary in the ease with which their referents can be identified such that, in some cases, no referent at all is automatically and uniquely identified. Failure to identify a unique referent might occur as the result of a number of factors, including the semantic and pragmatic content of the discourse and the speed required of comprehension processes by the speaker or reader. The possibility that pronouns sometimes fail to evoke unique referents has been discussed previously by Yule (1982), who points out that, in some discourse contexts, the identity of the entity referenced by an anaphor may be irrelevant to the reader or listener. Webber (1983) also suggests that, if there is no im-

mediate need to determine a unique referent, an anaphor may be left unresolved. Empirically, failure to resolve pronouns has been demonstrated by Greene et al. (1992). Their experiments investigated the difficulty of identifying a unique referent for a third person singular pronoun when two possible referents had been evoked in the discourse. Evidence for unique resolution was obtained only when reading rate was slow or readers could anticipate at exactly what point in the discourse the pronoun would occur. When reading rate was more normal (250 ms per word) or readers could not exactly anticipate the pronoun, the data suggested that no unique referent was identified.

The possibility that pronouns may sometimes be left unresolved complicates efforts to measure how difficult they are to comprehend. In particular, the time taken to read a pronoun (or the time to read a sentence containing a pronoun) is not an adequate measure. This is because reading times can reflect either time to successfully resolve a pronoun or time to process the pronoun but fail to resolve. One pronoun read in a given amount of time might be relatively easy to comprehend and so be identified with a unique referent, while another pronoun read in the same amount of time might be relatively difficult and left without a referent. In other words, reading time cannot be interpreted as a measure of comprehension difficulty unless it is combined with some measure of whether the pronoun was successfully resolved. Two methods have been typically adopted in previous research (cf. Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989; McKoon & Ratcliff, 1980b). One is to present the intended antecedent of the pronoun as a recognition test word at some point in the discourse after the pronoun. The reasoning that underlies this method is that successful resolution of the pronoun will increase the accessibility of its referent. This increase in accessibility will, in turn, facilitate the recognition decision about the referent when it

is presented as a test word. This method was used in Experiments 1 through 3. The second method, used in Experiments 4 through 6, is to use priming in word recognition to show that information given in the discourse with the pronoun is connected in memory to the referent, as it should be if the referent is correctly and completely understood (McKoon & Ratcliff, 1980b).

EXPERIMENT 1

Table 1 shows examples of the texts that were used in the experiment. Subjects read texts one line at a time, in a self-paced procedure. After the final line of a text, a single test word was presented for recognition (a decision as to whether or not the word had appeared in the text).

Table 1 also illustrates the design of the experiment: the accessibility of a discourse entity was manipulated pragmatically, by how closely it was related to the topic of its text, and syntactically, by using either the verb phrase or the compound construction. The referent entity (*deer* in Table 1) was introduced in the next to last sentence of its text, and it was the intended referent of the pronoun mentioned in the last sentence. It was also used as the test word that appeared after the final line of the text. The hypothesis was that the accessibility of the referent entity would be increased when it was more closely related to the topic and when it was introduced in a verb phrase. Increased accessibility was expected to result in faster reading time for the final sentence containing the pronoun, faster response time for the test word, or both.

Method

Subjects. Forty subjects participated in the experiment for credit in an introductory psychology class. Each subject participated in one 50-min session.

Materials. Twenty-four sets of four texts were written, each set with one critical referent noun. The four texts of a set implemented the variables of the experiment: the referent noun was used either in a com-

pound or in a verb phrase, and it was either more or less closely related to the topic of its text. The four texts of one set are shown in Table 1. For each of the four texts in a set, the next to last sentence stated the same information about the referent noun and a verb (e.g., *deer hunting* or *hunting deer*). The final sentences of the texts were the same in all four versions and referred to the referent noun with a pronoun (*He thinks they are really exciting to track*). The referent noun was stated only in the next to last sentence. The referent noun was also the test word for the experimental texts.

The mean lengths of both versions of the texts were 58 words, 5 sentences, and 7 lines as they appeared on a CRT screen. The last line of each text was always the entire final sentence of the text with no words from the preceding sentence.

There were 30 additional texts used as fillers in the experiment. These varied from 5 to 7 CRT lines in length, and averaged 50 words. Twenty of these had associated with them a single test word that did not appear in any of the filler or experimental texts. The test word for the other 10 was a word from the text. For each of these 30 texts, there was a true/false test sentence. Half of the test sentences were true and half false.

Procedure. All materials were presented to subjects on a CRT screen, and responses were made on the CRT's keyboard. Presentation and data collection were controlled by a real-time computer system.

The experimental session began with practice on 10 items presented one at a time for lexical decision. Subjects were instructed to respond to these items as quickly and accurately as possible, pressing the ?/ key on the keyboard if the test item was a word and the z key if it was not a word. These items were used to familiarize the subjects with the response keys.

After this practice, the experiment proper began. The texts were presented one at a time, with six of the fillers first, and then the remaining 24 fillers and the 24 experimental texts in random order. For each

text, first the instruction *Press space bar for next paragraph* appeared on the screen. When the subject pressed the space bar, there was a pause of 1000 ms, and then the first line of the text appeared. The line remained on the screen until the subject pressed the space bar again, and then the next line of the text appeared just below the first line. The subjects were instructed to press the space bar for the next line when they had read and understood the current line. The text continued in this way, with one additional line every time the space bar was pressed, until the last line of the text. When the space bar was pressed after reading of the last line, the screen was cleared and a test word appeared below where the last line had been. The test word was underlined by a row of asterisks. Subjects were instructed to respond *yes* (with the ?/ key) or *no* (with the z key) according to whether the test word had appeared in the preceding text. The test word remained on the screen until the subject pressed a response key, and then the screen was cleared. For the filler texts, the message *True-False Question* was then displayed, followed by the true/false question for the preceding text. Subjects answered the question by pressing the ?/ key for *true* and the z key for *false*. If the response was incorrect, the message *ERROR* was displayed for 2000 ms. After the true/false question, the next text began with the instruction to press the space bar.

Design. The two variables in the experiment were the topicality of the referent noun, and whether the noun was mentioned in a compound or a verb phrase. These two variables were crossed in a Latin square design, with four sets of materials (six per set) and four groups of subjects. Order of presentation of the texts was random, different for every second subject.

Results

For each text and each subject, means for the reading times of the texts' final sentences and means for response times to the

test words were calculated. Means of these means are presented in Table 2. Analyses of variance were performed on the means from the experimental design with both subjects and items as random variables; $p < 0.05$ was used unless otherwise noted.

First, the data for the test words are considered. For each text, the test word was the referent noun, the antecedent of the pronoun in the final sentence. If, for all four conditions, subjects interpreted the pronoun correctly during the time they were reading the final sentence, then response times to the test word should be equal across the conditions. The processes of interpreting the pronoun might be more or less difficult across conditions, but if the correct referent was always evoked by the pronoun then it should be equally accessible across conditions at the time the test word was presented. This is what the data show: there are no significant differences in response times to the test words (analyses of variance showed F's < 1.2). The standard error of the response times was 23.8 ms. Differences in error rates were also not significant, F's < 1.9.

Reading times show that there were differences in comprehension difficulty for the final sentences. It was hypothesized that interpretation of the pronoun would be difficult when the antecedent of the pronoun was in the modifier position in the compound. The data show this difficulty when the referent noun was low in topicality: reading times were longer when the noun was in a compound compared to when it was not. However, according to the discourse model theory, the difficulty should be reduced when the referent noun is more topical. This hypothesis was confirmed; increased topicality reduced reading times in the compound condition so that they were only slightly longer than in the verb phrase condition.

These effects were supported by analyses of variance. The main effect of compound versus verb phrase was significant, $F_1(1,39) = 10.2$ and $F_2(1,20) = 7.4$, as was the main effect of topicality, $F_1(1,39) = 21.8$ and $F_2(1,20) = 13.3$. The interaction of the two variables was marginally significant, $F_1(1,39) = 3.7$ and $F_2(1,20) = 4.3$. Planned tests showed that the difference between the compound and verb phrase conditions was significant when the referent noun was low in topicality, $F_1(1,39) = 11.2$ and $F_2(1,20) = 14.1$, but not when it was high in topicality, F's < 1.0. The standard error of the reading times was 52.5 ms.

For the true test questions, the mean response time was 2110 ms with 9% errors. For the false questions, the means were 2031 ms and 9% errors.

EXPERIMENTS 2 AND 3

Our interpretation of the results of Experiment 1 depends on the assumption that

TABLE 2
DATA FROM EXPERIMENT 1

Syntactic structure	Response times and error rates for test words			
	Low topicality text version		High topicality text version	
Compound	907 ms	5%	870 ms	2%
Verbal complement	893 ms	4%	886 ms	4%
Filler positive test words		1242 ms	21%	
Filler negative test words		1181 ms	15%	
Syntactic structure	Reading times for final sentences			
	Low topicality text version		High topicality text version	
Compound	2117 ms		1785 ms	
Verbal complement	1868 ms		1738 ms	

subjects understood the correct referents of the pronouns in the final sentences of the texts in all of the experimental conditions. This assumption is consistent with the finding that response times for the test words were equal across experimental conditions. However, the assumption might be wrong. An alternative possibility is that the pronouns were not understood at all, and that this is the reason that response times to the test words did not differ across the experimental conditions. By this alternative, the differences in reading times would represent differing degrees of *unsuccessful* efforts at understanding the final sentences, and there would be no way to determine whether the same pattern of reading times would hold for successful efforts. Experiments 2 and 3 were designed to rule out this alternative.

In both of these experiments, the same basic texts were used as in Experiment 1. However, there were two different possible final sentences. In one final sentence, the same pronoun referring to the critical referent noun was used as in Experiment 1 (*And he says they are really exciting to track* for the text in Table 1). In the second final sentence, a new noun was substituted for the pronoun (*And he says bears are really exciting to track*). This new noun had not been mentioned previously in the text.

In Experiment 2, the final sentence mentioned either the pronoun or the new noun, and following the final sentence, the referent noun was presented as a test word. If the pronoun in the pronoun version of the final sentence is understood as referring to the referent noun, and it is this processing that leads to the facilitation of response times when the referent noun appears as a test word, then response times should be facilitated only when the final sentence contains the pronoun, and not when it mentions the new noun. This was the prediction for the results of Experiment 2.

In Experiment 3, the two final sentences from Experiment 2, one with the pronoun and the other with the new noun, were used

and a new test word was introduced. The new test word was a "control" word picked from one of the earlier sentences of the text (e.g., *trips* for the texts in Table 1). There was also a second test word, the same referent noun test word as was used in the previous experiments. Again, we predicted response times to the test words from our assumption that the pronoun in the pronoun version of a final sentence is understood to refer to the referent noun. The pronoun version of the final sentence should facilitate response times for the referent noun test word relative to the new noun version, but response times for the control word should not be affected by which version of the final sentence is read.

Method

Subjects. For Experiment 2, there were 40 subjects and for Experiment 3, 24 subjects, all from the same population as in Experiment 1.

Materials. The basic texts from Experiment 1 were used in Experiments 2 and 3. For each text, a new final sentence was written. This sentence was almost the same as the old final sentence except that the pronoun was replaced by a noun. The new noun had not been mentioned previously in the text, but it plausibly fit the context of the text. There were slight changes in wording from the final sentences used in Experiment 1 to the sentences for Experiments 2 and 3, in order to keep both the pronoun and the new noun versions of the sentences about equally plausible. The mean length of the final sentences with pronouns was 8.4 words, and the mean length of the final sentences with new nouns was 8.9 words. For Experiment 2, the test word for each text was the critical referent noun (e.g., *deer*), the same as was used in Experiment 1. For Experiment 3, there were two possible test words, the referent noun and another control word that had appeared earlier in the text. For both experiments, the same filler paragraphs were used as in Experiment 1.

In these experiments, including all four

versions of the basic texts would have reduced power beyond acceptable limits. We compromised considerations of power with considerations of generality across versions by using two versions in Experiment 2, the high topicality compound version and the high topicality verb phrase version. In Experiment 3, only one version of the basic texts was used, the high topicality, compound version.

Method and design. The procedure was the same as that used in Experiment 1. For Experiment 2, there were two variables: whether the referent noun was stated in a compound or a verb phrase, and whether the final sentence contained the pronoun or the new noun. For Experiment 3, there were also two variables: the final sentence mentioned either the pronoun or the new noun, and the test word was either the referent noun or the control word. For both experiments, the two variables were combined in a Latin square design with four sets of materials and four groups of subjects. The order of presentation of the texts was random, different for every second subject.

Results

The data were analyzed as in Experiment 1, and are presented in Tables 3 and 4.

Experiment 2. When the final sentence contained the pronoun referring to the critical noun, the results of Experiment 2 replicated those of Experiment 1. Whether the critical noun was introduced in a verb phrase or a compound, high topicality should have made it easily accessible, and so, as is shown in Table 3, there should be little effect of syntactic structure on either response times for the referent nouns or reading times for the final sentences.

If processing of the pronoun in the final sentence facilitated responses to the critical noun test word, then replacing the pronoun in the final sentence with a new noun should slow responses to the test word. The data clearly show this effect.

Analyses of variance showed only one significant effect for response times for the referent nouns; when the final sentences contained the new nouns, response times were longer than when the sentences contained the pronouns, $F_1(1,39) = 18.1$ and $F_2(1,20) = 225.8$. The standard error was 27.1 ms. There were more errors on the test words when the final sentences contained the new nouns; these results were marginally significant with $F_1(1,39) = 3.7$ and $F_2(1,20) = 3.5$. There was also only one significant effect for reading times; reading times for the sentences with the new nouns were longer than reading times for the sentences with pronouns, $F_1(1,39) = 8.0$ and $F_2(1,20) = 5.2$. The standard error of the reading times was 102.2 ms. All other F's were less than 2.6.

For the true test questions, the mean response time was 1985 ms with 10% errors

TABLE 3
DATA FROM EXPERIMENT 2

Syntactic structure	Response times and error rates for test words				
	Pronoun final sentence			New noun final sentence	
Compound	948 ms	7%		1070 ms	8%
Verbal complement	926 ms	5%		1045 ms	10%
Filler positive test words		1263 ms	23%		
Filler negative test words		1150 ms	14%		

Syntactic structure	Reading times for final sentences	
	Pronoun final sentence	New noun final sentence
Compound	1961 ms	2199 ms
Verbal complement	2012 ms	2254 ms

TABLE 4
DATA FROM EXPERIMENT 3

| Test word | Response times and error rates for test words | | | |
	Pronoun final sentence		New noun final sentence	
Critical noun	884 ms	2%	1028 ms	10%
Control noun	1216 ms	17%	1219 ms	14%
Filler positive test words		1157 ms	23%	
Filler negative test words		1106 ms	8%	

| | Reading times for final sentences | |
	Pronoun final sentence	New noun final sentence
Compound, high topicality	1884 ms	1951 ms

and for the false questions, the means were 1941 ms and 14% errors.

Experiment 3. In Experiment 3, the final sentence contained either the new noun or the pronoun that was intended to refer to the referent noun. For the referent noun test word, responses should be facilitated only with the pronoun and not the new noun, as in Experiment 2, and the means in Table 4 show this facilitation. For the control test word, there should be no effect of whether the final sentence contained the pronoun or the new noun, and the data showed no effect.

Analyses of variance for response times to the test words showed a main effect for test word (referent noun or control word), $F_1(1,31) = 36.6$ and $F_2(1,23) = 147.8$, and a main effect of final sentence (pronoun or new noun), $F_1(1,31) = 4.8$ and $F_2(1,23) = 11.6$. The interaction of the two variables was significant, $F_1(1,31) = 4.2$ and $F_2(1,23) = 7.2$. Standard error for the response times was 26 ms. For error rates, the main effect of test word was significant, $F_1(1,31) = 17.9$ and $F_2(1,23) = 9.7$, as was the interaction of test word and final sentence, $F_1(1,31) = 6.7$ and $F_2(1,23) = 5.1$. The difference in reading times for the two versions of the final sentences was marginally significant, $F_1(1,32) = 15.2$ and $F_2(1,24) = 3.7$.

For true test statements, the mean response time was 1936 ms (8% errors), and for false test statements, 1941 ms (13% errors).

EXPERIMENTS 4, 5, AND 6

Experiments 1, 2, and 3 appear to show that the time required to comprehend a pronoun is a function of the accessibility of the pronoun's referent in the discourse structure. When accessibility is reduced, either via syntax, by introducing the referent with the compound rather than the verb phrase syntax, or via pragmatics, by making the referent less relevant to the discourse topic, then comprehension takes longer. This was shown in the reading times of the sentences containing the pronouns.

We pointed out that increased reading time does not by itself conclusively show that the pronouns were understood. In addition, some measure of the extent to which the pronouns were actually understood must be provided. Experiments 1, 2, and 3 used an immediate test of the antecedent of the pronoun (the referent noun) to provide evidence of comprehension. Immediate testing provides evidence about the relationships among discourse concepts that are available when both the discourse and the test item are in working memory at the same time (Corbett & Dosher, 1988; van Dijk & Kintsch, 1983; McKoon & Ratcliff, 1980b; 1986; 1989); in the present case, the relevant relationships are those among the pronoun, its intended referent in the discourse model, and the test word. From the results of Experiments 1, 2, and 3, we can conclude that those relationships were available to subjects at the time the test

word was presented. Whether understanding was complete, to the extent that the relationships among the pronoun, its intended referent, and information given in the discourse about the referent were all encoded into long term memory is still an open question (see McKoon & Ratcliff, 1989, for a case in which relationships available at immediate testing were not available at later testing). In Experiments 4, 5, and 6, we used a priming procedure to examine these relationships in long term memory.

The experiments involved a series of study test lists. For each list, subjects read four texts, and then they were given a list of test words for recognition (responding positively if a test word had appeared in one of the studied texts, and negatively if it had not). For the experimental texts, the test words of interest were the referent noun (e.g., *deer*) and a modifier from the final sentence (e.g., *exciting*). These two words were presented in immediately adjacent positions in the test list, with *exciting* following *deer,* and so they formed a "priming" pair. From previous research (McKoon & Ratcliff, 1980a; 1980b; Ratcliff & McKoon, 1978; Ratcliff & McKoon, 1988), it can be predicted that responses for the second word of the pair will be facilitated when they are closely related in memory by virtue of being from the same text (relative to being from different texts). The question is whether facilitation will be even further increased when the modifier *exciting* should be understood (by virtue of processing the pronoun) to describe the referent noun *deer.* Such further facilitation would be evidence that comprehension of the pronoun resulted in long-term memory encoding of the appropriate relationships between the referent and information given in the discourse about the pronoun.

In Experiment 4, the final sentence of a text contained either the pronoun for which the referent noun was the intended antecedent (. . . *they were exciting to track*), or the "new noun" of Experiments 2 and 3 (. . . *bears were exciting to track*). If sub-jects understand the final sentences completely, then *deer* should be more closely related in memory to *exciting* for the pronoun version of the final sentence than the new noun version, and this increased relatedness should lead to greater facilitation of responses to *exciting* by *deer* for the pronoun final sentence than the new noun final sentence. The results of the experiment followed this prediction.

In Experiment 4, only one version of each text was used, the high topicality, compound version. Experiments 5 and 6 were designed to check that the referent noun and the modifier were closely related in memory for both the high and low topicality versions of the text (Experiment 5) and for both the compound and verb phrase versions (Experiment 6).

Method

Materials. The same basic 24 texts were used as in Experiments 1, 2, and 3. The test words for these texts were the referent noun, the modifier from the final sentence, and two other words from the text. Thirty-two filler texts (30 of them the same as in Experiments 1, 2, and 3) each had four positive test words. Negative test words were chosen from a pool of 142 words that did not appear in any text.

Procedure. The experiments began with ten lexical decision test items, presented for practice on the response keys. This practice was followed by 14 study test lists. The first two study lists each contained four filler texts, and the remaining 12 each contained two experimental texts and two filler texts. The four study texts were presented in random order, one at a time, for 10 s for the filler texts and 11.5 s for the experimental texts. There was a 1.5-s blank interval between each text. After the four texts, a row of asterisks was presented for 1 s to signal that the test list was about to begin. The words in the test list were presented one a time. A word remained on the CRT screen until a response key was pressed (?/ for positive responses, z for negative re-

sponses). If the response was correct, then there was a blank screen for 200 ms, and then the next test word. If the response was incorrect, the word *ERROR* was displayed for 2 s. There was a total of 26 test words, 16 positive and 10 negative. After the 26th test word, two true/false test statements were presented, one at a time, with the ERROR message displayed for 2 s after incorrect responses. Then the next study test list began.

For study test lists containing experimental texts, the 16 positive test words were: the referent noun and the modifier from each experimental text, two other words from each experimental text, and four words from each filler text. A modifier was always tested later than the third position in the test list, and it was immediately preceded by the referent noun either from its own text or the other experimental text, depending on the experimental condition. The other words from an experimental text were tested later in the test list than the modifier. Otherwise, the order of the test words was random. No word appeared more than once in a test list.

Design. In all three experiments, the first variable was whether the modifier was preceded in the test list by the referent noun from its own text or from the other experimental text. In Experiment 4, the second variable was whether the final sentence of a text was studied in the pronoun version or the new noun version, and only the high topicality, compound versions of the texts were used. In Experiment 5, the second variable was whether the context was high topicality or low topicality. The final sentence was always the pronoun version, and the referent noun always appeared in a compound. In Experiment 6, the second variable was whether the referent noun was presented in a compound or a verb phrase. The final sentence was the pronoun version, and the high topicality texts were used. In each experiment, the two variables were crossed in a Latin square design, with four groups of subjects and four sets of texts. There were 52 subjects in Experiment 4, 32 in Experiment 5, and 24 in Experiment 6.

Results

In Experiment 4, the referent noun should be more closely related in memory to the modifier when the final sentence referred to the referent noun with a pronoun than when it did not. Thus, in the test list, the referent noun should facilitate responses for the modifier more when the final sentence referred to the referent noun. This is the pattern shown in Table 5. With the pronoun in the final sentence, response times to the modifier are facilitated 160 ms when the referent noun comes from the same text relative to a different text. With the new noun in the final sentence, the facilitation is only 53 ms. This interaction was significant, $F_1(1,51) = 9.5$ and $F_2(1,23) = 6.0$. The main effect of same versus different text prime was also significant, $F_1(1,51) = 27.7$ and $F_2(1,23) = 33.9$. Which version of the final sentence was used had no significant effect, F's < 1.0. The standard error of the response time means was 19 ms.

TABLE 5
Data from Experiment 4

	Response time and error rates for test words				
Prime	Pronoun final sentence			New noun final sentence	
Critical noun from same text	714 ms	9%		764 ms	6%
Critical noun from different text	874 ms	15%		817 ms	19%
Filler positive test words		784 ms	12%		
Filler negative test words		944 ms	25%		

Part of the significant interaction-effect (but only part) comes from the pattern of response times in the conditions for which the prime comes from a different text than the modifier; responses are slower when the final sentence contains a pronoun than when it contains the new noun. This difference has no obvious explanation. For errors, the main effect of same versus different text for the prime was significant, $F_1(1,51) = 30.0$ and $F_2(1,23) = 14.9$. Other F's for error rates were less than 1.8.

True test statements averaged 2079 ms in response time, and 17% errors, and false statements, 2077 ms and 12% errors.

In Experiments 5 and 6, the hypothesis was that the relation between the pronoun in the final sentence and the referent noun is encoded in memory equally well, whether the text is presented in the high topicality or low topicality versions, or whether the referent noun is presented in a compound or a verb phrase. As a result, there should be equal amounts of facilitation from the referent noun to the modifier in all cases. The results in Tables 6 and 7 confirm this prediction.

In Experiment 5, there is about the same amount of facilitation with the high topicality texts (55 ms) as with the low topicality texts (48 ms). Overall, the subjects in Experiment 5 were faster than those in Experiment 4 (see response times for filler test items), so the facilitation is somewhat reduced in size. The main effect of whether the prime comes from the same or a different text than the modifier is significant, $F_1(1,31) = 9.9$ and $F_2(1,23) = 7.1$. The interaction with text version did not approach

significance, F's < 1.0. The main effect of text version approached significance in the subjects analysis, $F_1(1,31) = 3.5$, but was less than one in the items analysis. The standard error of the response time means was 19 ms. For errors, the main effect of same versus different text for the referent noun was significant with the subjects analysis, $F_1(1,31) = 4.1$, but not with the items analysis, $F_2(1,23) = 2.1$.

True test statements averaged 2071 ms in response time, and 15% errors, and false statements, 1990 ms and 13% errors.

The results of Experiment 6 (Table 7) show that the amount of facilitation is not significantly affected by whether the referent noun appeared in a compound (49 ms of facilitation) or a verb phrase (64 ms). The main effect of same versus different text for the prime was significant, $F_1(1,23) = 10.0$ and $F_2(1,23) = 9.8$. All other F's were less than 1.0. The standard error of the response time means was 20 ms. Same versus different text for the prime also significantly affected error rates, $F_1(1,23) = 31.6$ and $F_2(1,23) = 11.8$. Again, no other F's were greater than one.

True test statements averaged 2126 ms in response time, and 15% errors, and false statements, 2007 ms and 11% errors.

Summary. Experiments 4, 5, and 6 used a priming procedure to examine the long term memory representation of the relations between the referent entity (e.g., *deer*) and information given in the text about that entity. In the final sentence, the information that *they are exciting to track* should be understood such that *exciting* is encoded into long term memory as describing *deer*.

TABLE 6
DATA FROM EXPERIMENT 5

Prime	Response time and error rates for test words			
	High topicality text version		Low topicality text version	
Critical noun from same text	665 ms	8%	691 ms	5%
Critical noun from different text	720 ms	10%	739 ms	10%
Filler positive test words	714 ms	11%		
Filler negative test words	855 ms	26%		

TABLE 7
DATA FROM EXPERIMENT 6

Prime	Compound		Verbal complement	
	Response time and error rates for test words, syntactic structure			
Critical noun from same text	678 ms	1%	671 ms	4%
Critical noun from different text	727 ms	12%	735 ms	13%
Filler positive test words		705 ms	10%	
Filler negative test words		879 ms	31%	

If so, then a response to *deer* in the test list should facilitate a response to *exciting*, more so than if the sentence had said that *bears are exciting to track*. Experiment 4 demonstrated this result, and Experiments 5 and 6 showed that the same result obtained whether *deer* was more or less topical and whether it was introduced in a verb phrase or a compound.

GENERAL DISCUSSION

A discourse model is the representation of information that is built during comprehension of a text or discourse. As comprehension proceeds through a text, the discourse model is continually updated and revised to include new input and to reflect the impact of new input on earlier information. In the discourse model theory assumed as the background for the experiments in this article, the model is made up of the entities evoked by linguistic and contextual information, the relations among the entities, and their accessibilities relative to potential referential cues.

The discourse model that we assume differs from previous psycholinguistic approaches in two ways. First, we propose that the accessibility of a discourse entity is a function of a number of factors, both linguistic and nonlinguistic, arising from explicit information in the text as well as from contextual information, pragmatic knowledge, and speaker/writer and listener/reader goals. In addition, the accessibility of an entity for later reference is determined by the cue with which it is referenced. A given entity may be quite accessible from one cue, but relatively inaccessible from another. Thus, accessibility is an interaction between entities in the discourse model and the cues used by the speaker/writer to evoke those entities.

The experiments presented in this article support the discourse model view by showing that both the morphosyntactic and the pragmatic context in which an entity is introduced into a discourse determine its accessibility for later reference. In Experiment 1, a referent entity (*deer*) was introduced in a morphosyntactic context that made it either more accessible (a verb phrase, *hunting deer*) or less accessible (a compound, *deer hunting*). Reading times for a sentence containing a pronominal anaphor for the referent entity were correspondingly faster when the entity had appeared in a verb phrase versus compound. The referent was also introduced in two pragmatic texts; in one case, it was more closely related to the topic of its discourse than the other. Again, reading times for the sentence with the pronoun reflected accessibility, with faster reading times when the referent was more topical. In fact, when the referent entity was highly related to the discourse topic, reference in the compound condition was not significantly more difficult than reference in the verb phrase condition.

These results validate the claim that short term memory for text comprehension contains a representation of the relative accessibilities of discourse entities, accessibilities that are jointly determined by pragmatic and syntactic factors. The results also support the claim that naturally occurring examples of antecedents in compounds

(e.g., *Kal Kan* in the compound *Kal Kan cat*) are grammatically well formed and that they are neither "performance errors" nor the result of some type of pragmatic salvaging of otherwise ungrammatical constructions (cf. Ward et al., 1991). The fact that such examples are frequently produced in natural discourse does not necessarily entail that they are understood by the hearer/reader. But the psycholinguistic data presented in this article indicates that they are and that they are subject to the same types of pragmatic variables as are other kinds of anaphora. The pragmatic variable in our experiments, topicality, affected reference for both compound and noncompound constructions. Furthermore, placing the word that evokes the referent entity in a compound internal position reduced the accessibility of that entity, just as a modifier position reduces the accessibility of other entities (McKoon et al., in preparation; Rothkopf et al., 1986; Rothkopf et al., 1988).

The results from the six experiments in this article, taken as a whole, also demonstrate the importance of using converging kinds of experimental data. It would not be possible for us to support our conclusions from measurements of sentence reading times alone. For example, we found that reading times were slowed when the referent entity for the pronoun in the final sentence was introduced within a compound. But we could have found that reading times in this condition were quite fast; this could have happened if the pronoun were uninterpretable and subjects quickly realized that it was uninterpretable. In this instance, the reading time data would have seemed to counter our hypotheses. However, an uninterpretable pronoun would have led to slow response times when the referent word was tested immediately after the final sentence, allowing us to correct what would have been erroneous conclusion.

Likewise, it would not be possible, with reading times and immediate testing alone, to conclude that a pronoun was completely

understood such that all the relevant relationships among the referent entity and information in the discourse were encoded into long term memory. The delayed testing priming results are required for that conclusion. Thus, only by simultaneous consideration of the sentence reading times, the test word response times, and the priming results can our interpretations of sentence reading times be fully jusitified.

Through these converging sets of data, we argue that the difficulty of comprehension for a pronoun depends on the accessibility of the discourse entity to which the pronoun is being used to refer. Pronoun comprehension is not viewed as a process that depends on the pronoun alone or even primarily. The issue for the comprehension system is not how to use a pronoun to access the intended referent. Instead, the issue is how the discourse model is constructed from the discourse in such a way that pronouns can be automatically and correctly interpreted.

REFERENCES

BRANSFORD, J. D., BARCLAY, J. R., & FRANKS, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 3, 193–209.

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

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

CORBETT, A. T., & CHANG, F. R. (1983). Pronoun disambiguation: Accessing potential antecedents. *Memory and Cognition*, 11, 283–294.

CORBETT, A. T., & DOSHER, B. A. (1978). Instrument inferences in sentence encoding. *Journal of Verbal Learning and Verbal Behavior*, 17, 479–491.

FODOR, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, 4, 155–209.

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

GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1–67.

GREENE, S., McKOON, G., & RATCLIFF, R. (1992). Pronoun resolution and discourse models. *Journal*

of Experimental Psychology: Learning, Memory and Cognition, **18,** 266–283.

GROSZ, B. (1977). The representation and use of focus in a system for understanding dialogs. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence,* Los Altos: William Kaufmann, Inc.

GROSZ, B., & SIDNER, C. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics,* **12,** 175–204.

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

HEIM, I. (1982). *The semantics of definite and indefinite noun phrases.* Unpublished Ph.D. dissertation, University of Massachusetts at Amherst.

HINTZMAN, D. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review,* **95,** 528–551.

JARVELLA, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior,* **10,** 409–416.

JOHNSON-LAIRD, P. (1983). *Mental Models.* Cambridge, MA: Harvard University Press.

KAMP, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof (Eds.), *Formal methods in the study of language. Part 1* (pp. 277–322). Amsterdam: Mathematical Centre Tracts.

KARTTUNEN, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics VII: Notes from the linguistic underground* (pp. 363–386). New York: Academic Press.

KINTSCH, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

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

KINTSCH, W., & VAN DIJK, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review,* **85,** 363–394.

KINTSCH, W., & KEENAN, J. M. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology,* **5,** 257–274.

LAKOFF, G., & ROSS, J. (1972). A note on anaphoric islands and causatives. *Linguistic Inquiry,* **3,** 121–125.

LEVI, J. N. (1978). *The syntax and semantics of complex nominals.* New York: Academic Press.

LIEBER, R. (1984). Grammatical rules and sublexical elements. In *Papers from the 20th parasession on lexical semantics, Chicago Linguistic Society* (pp. 187–199).

MALT, B. (1985). The role of discourse structure in understanding anaphora. *Journal of Memory and Language,* **24,** 271–289.

MATHEWS, A., & CHODOROW, M. (1988). Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location and depth of embedding. *Journal of Memory of Language,* **27,** 245–260.

MATTHEWS, P. (1974). *Morphology.* Cambridge, UK: Cambridge University Press.

MCKOON, G. (1977). Organization of information in text memory. *Journal of Verbal Learning and Verbal Behavior,* **16,** 247–260.

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

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

MCKOON, G., & RATCLIFF, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **12,** 82–91.

MCKOON, G., & RATCLIFF, R. (1989). Assessing the occurrence of elaborative inference with recognition: Compatibility checking vs compound cue theory. *Journal of Memory and Language,* **28,** 547–563.

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

MOHANAN, K. P. (1986). *The theory of lexical phonology.* Dordrecht: Reidel.

MORGAN, J. (1978). Toward a rational model of discourse comprehension. In D. Waltz (Ed.), *Proceedings of TINLAP-2* (pp. 109–114). New York: Association for Computational Machinery.

MORROW, D., BOWER, G., & GREENSPAN, S. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language,* **28,** 292–312.

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

MURPHY, G. L. (1985). Psychological explanations of deep and surface anaphora. *Journal of Pragmatics,* **9,** 171–198.

NICOL, J., & SWINNEY, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research,* **18,** 5–20.

OAKHILL, J., GARNHAM, A., & VONK, W. (1989). The on-line construction of discourse models. *Language and Cognitive Processes,* **4,** 263–286.

POSTAL, P. (1969). Anaphoric islands. In R. Binnick et al. (Eds.), *Papers from the fifth regional meeting of the Chicago Linguistic Society* (pp. 205–255).

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

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.

RATCLIFF, R., & McKOON, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, 95, 385–408.

ROTHKOPF, E., BIESENBACH, B., & BILLINGTON, M. (1986). *Syntax violations as error feedback during rapid reading: Suggestions for a new readability measure supplement* (Technical memorandum). AT&T Bell Laboratories.

ROTHKOPF, E., KOETHER, M., & BILLINGTON, M. (1988). *Why are certain sentence constructions mnemonically robust for modifiers?* (Technical memorandum). AT&T Bell Laboratories.

SANFORD, A. J., & GARROD, S. C. (1981). *Understanding written language*. New York: Wiley.

SIDNER, C. (1979). *Towards a computational theory of definite anaphora comprehension in English discourse*. Cambridge, MA: MIT dissertation.

SIDNER, C. (1981). Focusing for interpretation of pronouns. *American Journal of Computational Linguistics*, 7.4, 217–231.

SIMPSON, J. (1983). *Aspects of Warlpiri morphology and syntax*. Cambridge, MA: MIT dissertation.

TANENHAUS, M. K., & CARLSON, G. N. (1990). Comprehension of deep and surface verb phrase ana-

phors. *Language and Cognitive Processes*, 5, 257–280.

VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

WARD, G. L. (1985). *The semantics and pragmatics of preposing*. Philadelphia: University of Pennsylvania dissertation. Reprinted in 1988, New York: Garland.

WARD, G., SPROAT, R., & McKOON, G. (1991). A pragmatic analysis of so-called anaphoric islands. *Language*, 67, 439–474.

WEBBER, B. (1979). *Elements of discourse understanding*. New York: Garland Press.

WEBBER, B. (1983). So what can we talk about now? In M. Brady and R. Berwick (Eds.), *Computational Models of Discourse* (pp. 331–371). Cambridge, MA: MIT Press.

WILSON, D., & SPERBER, D. (1979). Ordered entailments: An alternative to presuppositional theories. *Syntax and Semantics*, 11, 299–323.

YULE, G. (1982). Interpreting anaphora without identifying reference. *Journal of Semantics*, 1, 315–322.