

University of Massachusetts Occasional Papers in Linguistics

Volume 12 *University of Massachusetts
Occasional Papers in Linguistics - Special
Issue: Psycholinguistics- Volume 9*

Article 6

1986

Anaphora and Discourse Structure

Barbara C. Malt

University of Massachusetts, Amherst

Follow this and additional works at: <https://scholarworks.umass.edu/umop>



Part of the [Psycholinguistics and Neurolinguistics Commons](#)

Recommended Citation

Malt, Barbara C. (1986) "Anaphora and Discourse Structure," *University of Massachusetts Occasional Papers in Linguistics*: Vol. 12 , Article 6.

Available at: <https://scholarworks.umass.edu/umop/vol12/iss3/6>

This Article is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in University of Massachusetts Occasional Papers in Linguistics by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

ANAPHORA AND DISCOURSE STRUCTURE

BARBARA C. MALT

UNIVERSITY OF MASSACHUSETTS, AMHERST

The interpretation of an utterance often depends on information contained in earlier utterances.¹ By itself, the utterance "Yes, he did" indicates little more than that someone male is being referred to, and that he has done something. When "Yes, he did" is preceded by the question, "Did John ever return your book?", then "he" can be taken to refer to John, and the action he has done can be understood to be returning the book. To fully understand an utterance, then, a listener or reader must integrate information in one utterance with information in another, and understanding how people do so is an important part of understanding how discourse is understood.

Both "he" and the missing verb phrase after "did" are instances of anaphors, devices that refer back to previously mentioned words or concepts in a discourse. Much work has addressed the question of how anaphors are understood, and a large number of factors have been identified that play a role in the comprehension process. Gender, number, and animacy of course influence the choice of referent for pronouns, and syntactic constraints may affect the choice as well (see Shwartz, 1981; Springston, 1975). The comprehension process is also influenced by semantic properties of the verbs and connectives in the sentences (Caramazza, Grober, Garvey, & Yates, 1977; Ehrlich, 1980; Shwartz, 1981; Springston, 1975), and by the distinction between given and

new and focussed and non-focussed information (Carpenter & Just, 1977; Yekovitch, Walker, & Blackman, 1979).

A simple model accommodates the role of such factors in assigning meaning to an anaphor: When the anaphor is encountered, a backward search takes place for something that will provide an interpretation for it. The search is through the contents of short-term memory, which will generally be a representation of the most recent sentence or two preceding the one containing the anaphor. If there is more than one potential antecedent, the choice of interpretation is made using some combination of the various cues mentioned above.

A model like this has generally been either implicit or explicit in most work on understanding anaphors, and indeed, for subjects reading experimental sequences of simple declarative sentences, it is a plausible model. In understanding natural discourse, however, integration occurs at many different levels besides the resolution of anaphora. Taken as a complete model, the simple backward search model assumes that anaphora resolution is carried out using only basic syntactic, semantic, and pragmatic analyses of individual sentences. Such an assumption in turn implies either that the output of other, higher-level discourse processes is not available to the anaphor interpretation process, or that the output has no bearing on the problem of interpreting an anaphor.

The alternative possibility is that higher-level discourse processes do indeed influence anaphora resolution. In particular, in natural discourse, individual utterances vary both in the kind of information they convey, and in the role that that information plays in relation to nearby utterances and the discourse as a whole. For example, an utterance may be a question, or an assertion, or a demand; and it may be part of the main topic, or a digression, or an interruption. At some point in the integration process, a reader must understand not only how expressions such as pronouns relate to other expressions such as noun phrases, but also how the complete utterances relate to one another and to the larger discourse. If understanding these latter relations is a central and ongoing part of discourse comprehension, then the process of resolving anaphora, which requires location of information that bears a certain relation to the anaphor, may not be independent of that process. The possibility and plausibility of anaphoric reference may be influenced by such variations in utterance relations, and the listener or reader may be able to use this fact in interpreting an anaphor.

Some suggestion in this direction is made by Kintsch and van Dijk (1978). They propose that retention of particular information in short-term memory is influenced by its hierarchical

level of importance to a discourse as a whole. According to their model, important propositions are retained longer than less important ones and hence will be more readily available for integration with incoming information, including anaphors. Thus, to the extent that an anaphor is likely to depend on high-level information for its interpretation, the recognition of the importance of a proposition to a discourse may affect how easily a subsequent anaphor can be understood.

It is not clear, though, that the role of information in a discourse as a whole is the most useful level for considering the influence of utterance relations on anaphor resolution. A variety of evidence suggests that anaphoric reference is a very local process. Marslen-Wilson, Levy, and Tyler (1982), using a hierarchical analysis of a story, found that pronouns and ellipsis occurred almost exclusively within the same lowest-level event unit of a story. Grosz (1977) found that pronominal reference generally occurred within a particular sub-task of a larger task-oriented dialogue. Pronouns are understood most readily when the antecedent is in an immediately preceding clause, and the time to understand increases as a function of the distance of the antecedent from the anaphor (Carpenter & Just, 1977; Clark & Sengul, 1979; Ehrlich, in press). Furthermore, if the focus of a series of utterances does not remain on a particular entity, anaphoric reference to that entity rapidly becomes difficult to interpret (Garrod & Sanford, 1982; Lesgold, Roth, & Curtis, 1979). These findings suggest that recent information is the most relevant for interpreting anaphors, even if somewhat more distant information has greater importance to the discourse as a whole.

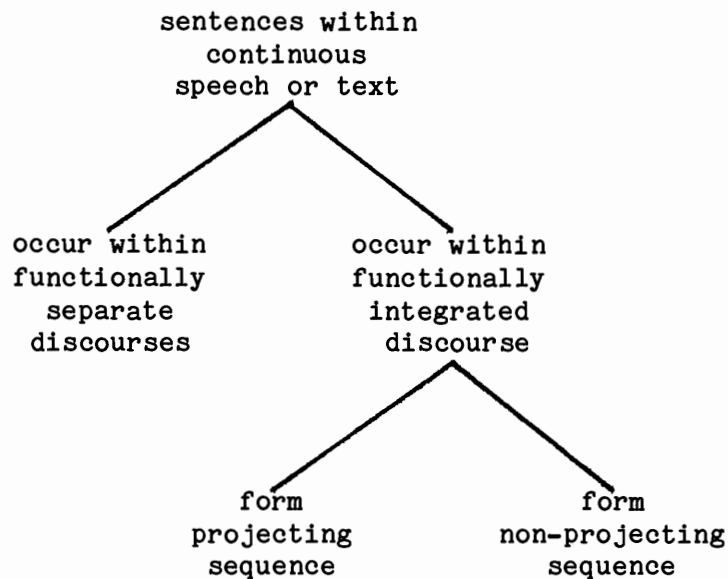
There is in fact a great deal of structure in local segments of discourse. Although discourse structure at this level has received little attention in the psychological literature, various aspects have been discussed in some detail by linguists and can provide a starting point for psychological considerations. The next section will draw on a number of linguistic analyses to outline of some of the local, between-utterance structure that exists in discourse. It will also suggest how the occurrence of anaphoric reference may be affected by such structure. Once this background has been provided, four experiments will be presented that test whether there is an influence of discourse structure on understanding anaphora, and if so, what that influence might be.

Relations between utterances

Relations between utterances may be divided according to several major distinctions, as shown in Figure 1.

Figure 1

Major Divisions in Between-Utterance Relations



The first division is between utterances that are part of the same functional piece of discourse and utterances that are not. Most often, utterances that occur near one another within a single continuous piece of speech or writing are part of the same discourse. The relations they may then bear to one another are described by the right branch of the diagram, which will be taken up later. Within a single continuous piece of writing or speech, however, there are sometimes portions that, although embedded in the same physical sequence, do not form part of the same discourse per se; they function as more or less independent pieces of discourse. This is the left branch of the diagram, which will be discussed first.

Functionally separated discourse

Functionally distinct pieces of discourse embedded in a single physical sequence can come about in a number of ways. One prime example is what Jefferson (1972) has termed a "side

sequence." A side sequence is a temporary break in an ongoing conversation. A topic other than the main topic is discussed during the break, and the main conversation is then resumed.

One common type of side sequence is an interruption, as in the following example:

Banker: Have you made a decision about the property?
 Client: Yes, I believe my company would benefit from
 acquiring Boardwalk.
 Waiter: Would you care to order now?
 Banker [to waiter]: No, thank you, not just yet.
 Banker [to client]: Are you prepared to sacrifice your
 hotels in order to acquire it?

This sequence illustrates how a side sequence within a conversation may influence the distribution of reference. The exchange about ordering, which constitutes the side sequence, does not contain references to anything said in the previous part of the conversation, and the final sentence about hotels not only is not tied to the ordering exchange, but refers instead to the original topic. This pattern of reference is typical of a portion of discourse that contains a side sequence: An utterance within the side sequence, by definition, is unlikely to refer to an element from the main part of the conversation, and once the intervening segment is finished, a subsequent anaphor is unlikely to refer back to information in that segment.

Another type of functional separation is that between dialogue and narrative within, for example, a short story or novel. Though dialogue and narrative together move the reader through a single story, each has its own integrity. Since narration is purely for the benefit of the reader and not part of the spoken exchange, the participants of a reported dialogue are generally unaware of narrative remarks. The narrative may even provide information for the reader about situations or motivations unknown to the characters as the dialogue takes place. The two thus cannot freely cross-refer, as in the following example:

Frank and Brian were walking down the street.
 "These elm trees look sick," Frank remarked.
 He pointed to the peeling bark.
 "I did, too," Brian replied.

Brian cannot anaphorically refer back to the sentence mentioning pointing because that sentence was not part of the dialogue between the two speakers. Again, this functional separation results in constraints on where anaphoric reference may occur. An anaphor mentioned within the dialogue will usually have its antecedent within the dialogue even if the preceding utterance was

narrative, and an anaphor occurring in narrative likewise will usually have its antecedent in earlier narrative.

Thus, within a single piece of writing or speech, some portions of a discourse may be functionally separate from other portions, and anaphoric reference will be constrained by this separation. Utterances within the different sub-discourses will be relatively independent of one another, even if they occur in close or adjacent positions, with the result that anaphors are most likely to refer back to information within the same functional piece of discourse, even if segments of another functional discourse intervene.

Functionally integrated discourse

Projecting sequences. Within any single functional portion of discourse, utterances will bear certain relations to one another, and these relations constitute the right-hand branch of Figure 1. There are two major divisions in within-discourse utterance relations, the first of which can be termed "projecting sequences." Utterances that form projecting sequences are predictable sequences following certain well-established conventions. Because the sequences are predictable ones, they have the characteristic that occurrence of the first utterance may project an expectation of the second utterance.

Some of the sequences that fall under this category are those that Schegloff and Sacks (1973) call "adjacency pairs." Adjacency pairs are two-sentence sequences such as question-answer, greeting-greeting, and offer-acceptance/refusal pairs, as in the following examples:

Did you see Susan today?
 -- Yes, I did.
 Hello.
 -- Hi.
 Would you like some coffee?
 -- No, thanks.

The important characteristic of these pairs for the present purposes is the limitation of possible second utterances that is imposed by the first utterance. An offer, for example, solicits a response that is an acceptance or refusal; questions require answers; and greetings require return greetings. Although Schegloff and Sacks define adjacency pairs as utterances spoken by two different people, other types of sequences also exhibit predictability of the second utterance from the first. An utterance beginning "On the one hand..." projects a closely following utterance with contradictory content, usually beginning "On the other hand..." Expressions such as "in the first

place...in the second place..." or "first of all...second of all..." are also examples of projecting sequences. Despite Schegloff and Sacks' term "adjacency pair," all these projecting sequences may be separated from each other by intervening sentences; the expectation of the second utterance can be retained over some extent of separation.

Just as anaphoric reference is most likely to occur within a single functional discourse, so it also is more likely to occur between two members of an adjacency pair than between one member of the pair and other nearby discourse. By definition, the second member of a projecting sequence is a remark or response that relates to the first member of the pair. Thus, the following sequence sounds natural,

On the one hand, Ray does want to come with us.
Heather told him the place is nice.
On the other hand, he knows he should be studying.

but in contrast, this sequence sounds odd:

On the one hand, Ray does want to come with us.
Heather told him the place is nice.
On the other hand, she knows she should be studying.

This latter sequence is odd because the pronoun "she" makes reference not to someone mentioned in the first half of the projecting sequence, but instead to someone mentioned in the intervening sentence.

An answer to a question is also more likely to make anaphoric reference to elements in the first half of the projecting sequence -- the question -- than to elements in other utterances in the discourse. If a response does not refer at least indirectly to the question, then it probably will not be taken to be an adequate answer. Thus, following this invitation,

"Would you like to come to Point Reyes with us?
The weather should be good."

the response "Yes, I would" is more appropriate and hence more likely than the response, "Yes, it should be." Here again, the presence of a particular type of sequence in a segment of discourse allows certain predictions to be made about where anaphoric relations will exist.

Non-projecting sequences. The remaining division in Figure 1 is that labelled "non-projecting sequences." Many utterances in a text or conversation, of course, are not part of highly constrained sequences, and they constitute the category of

non-projecting sequences. In non-projecting sequences, a given utterance may potentially be followed by an utterance having any one of a large number of relations to it. Names and descriptions of some of the possible relations have been compiled by various investigators (e.g. Clark, 1977; Hobbs, 1979; Halliday & Hasan, 1976; Longacre, 1979). Halliday and Hasan (1976) divide possible relations into additive, adversative, causal, and temporal, each of which can further be broken down into sub-categories. An additive relation, for example, exists between utterances such as

"I saw him do it with my own eyes."
 "Furthermore, John saw him do it, too."

and a causal relation exists between utterances such as,

"John slept through his alarm."
 "Consequently, he was in a bad mood all morning."

The exact names and divisions of other possible relationships are not critical for our purposes here; the point is simply that such relations exist between utterances in a discourse.

Many of these relations are signalled by particular words or phrases. For instance, words such as "yet," "only," "but," and "however" beginning an utterance signal a contradiction of an earlier utterance. Others such as "therefore," "for this reason," and "as a result" signal a causal relation. In contrast to the projecting sequences, though, the relations cannot be determined until the second utterance is given; the first utterance does not limit what sort of relation the second will hold to it. There is evidence that meaning is more available in short-term memory for clauses expressing certain kinds of causal-temporal relations than for others, a difference attributed to the differing dependence of later clauses on the earlier clause for interpretation (see Townsend, 1983; Townsend & Bever, 1978). In general, however, the dependence of any one utterance on another is much less strictly determined than in sequences involving projecting pairs or functional separations. As a result, the likelihood of anaphoric reference is more variable, and the relative probability that one sentence or clause will contain the antecedent to an anaphor in another must be determined by a more complete analysis of the meaning of the sentence in order to determine causal-temporal relations.

The influence of structure on interpreting an anaphor

Different relations between utterances thus do appear to have different consequences for the occurrence of anaphoric reference.

We are now in a position to consider how these relations might affect the on-line process of interpreting an anaphor.

A prime possibility is that differences in utterance relations suggest differences in relative utterance importance. Discourse structure at more global levels has often been discussed in terms of variations in importance. Stories have been described as hierarchically structured, where high-level propositions are more important than low-level ones (e.g. Kintsch & van Dijk, 1978; Mandler & Johnson, 1977; Rumelhart, 1975; Thorndyke, 1977), and individual paragraphs and smaller segments within paragraphs also seem to be composed of central information supported by less important information (e.g. Hinds, 1979; Longacre, 1979). One major processing benefit of recognizing these variations in importance may be in overcoming the limits of short-term memory: if a reader or listener can identify certain information as central to understanding a discourse, then that information can be selectively retained in short-term memory for continued use. Kintsch and van Dijk's (1978) proposal that high-level propositions from a text hierarchy are held through extra processing cycles makes just such an assumption.

Many of the utterance variations described in the preceding pages may also provide information about importance, but at a more local level. Dialogue may be of greater importance than surrounding narrative; utterances about the main topic of conversation should be more important than those in a side sequence; statements of fact may be more prominent than evaluative or explanatory utterances following them; and the first half of a projecting sequence may be more important than non-projecting utterances nearby. If readers do selectively retain information on the basis of its importance, then these local variations in importance may play a role in that selection. Dialogue may be retained over narrative, main topic information over digressions, and so on. Such selection would have a direct consequence for understanding anaphora: an anaphor would be interpreted more easily when it refers to this locally important information than when it refers to information less likely to be retained in short-term memory. Stated more generally, this hypothesis suggests that discourse structure will affect understanding an anaphor by influencing availability of the information needed for its interpretation. It is this possibility that will be the focus of the experiments to be presented.

One alternative explanation for discourse effects might be constructed in terms of search cues. The representation of information in the short-term buffer may preserve explicit cues to relatedness that are present in the text. The distinction between descriptive sections and pieces of dialogue, for instance, may also exist in the memory representation. Likewise, the

distinctions between a question-answer sequence and a declarative sequence in which it might be embedded, or between a side sequence and a larger conversation, might also exist in memory. Since the antecedent for an anaphor occurring in dialogue is likely to be within other dialogue, the antecedent for a response to a question in the question itself, and so on, these distinctions may provide a cue to the backward search procedure about where missing information can be found. Understanding an anaphor would then be easier when such a cue is present than when one is not. According to this explanation, discourse structure does not differ from factors such as the gender or number of a pronoun in the role that it plays in comprehension. The sufficiency of this alternative explanation will also be examined in the experiments to be presented.²

Experiments

These experiments examine the influence of discourse structure on the process of understanding anaphors. The first concern is whether anaphor resolution is in fact influenced by relations between utterances. All four experiments to be presented provide evidence that discourse structure does have an effect on anaphor resolution. The second concern is whether relations between utterances indicate what information should be held for integration with future utterances, or whether they have the more restricted function of providing additional search cues for locating information. Experiments 1 and 2 demonstrate one type of between-utterance relation that does influence ability to interpret an anaphor, and also demonstrate another that does not. These two results together suggest an answer in favor of the importance hypothesis. Experiments 3 and 4 further test this suggestion and refine its interpretation by exploring the effects of other manipulations.

These experiments use one particular type of anaphor, Verb Phrase Ellipsis (VPE). VPE exists when a sentence contains a dangling auxiliary or infinitive marker (to), and the verb phrase that would otherwise follow is omitted (see Hankamer & Sag, 1976; Sag, 1979). Thus, sentences such as "Yes, I did," "No, she wasn't," and "Susan didn't want to" are instances of VPE. One advantage of using VPE is that missing verb phrases cannot usually be interpreted using only non-linguistic context, as many anaphors (such as pronouns) can (Hankamer & Sag, 1976). To interpret VPE, a reader or listener must use information from a preceding sentence. This fact allows more certainty that subjects are actually locating and using information from the presented sequence when they interpret the anaphors. VPE is also generally difficult to understand when separated from the antecedent by any

distance (perhaps because the anaphor itself provides little information about the antecedent; see Murphy, 1982), so its use allows ample room for the costs and benefits of various manipulations to be seen.

Experiment 1

Experiment 1 was designed as a first test of the hypothesis that discourse structure influences ease of interpreting an elliptical utterance. There were two manipulations in this experiment. One used utterances from functionally separate discourses, and the other used utterances in projecting pairs. For both, there are explicit cues to where relatedness will be found: it will be found between utterances forming part of the same functional discourse in the first case, and between members of the projecting sequence in the second case. Any effect of between-utterance relations should be seen most readily in these two cases, where the cues are strongest. If both of these manipulations influence the resolution process, then it will be clear that diverse sorts of relations can provide cues for resolving anaphora. If one but not the other type of manipulation shows an effect, then the characteristics of that class of between-utterance relations will suggest some limitations on what the effect of structure may be.

All target sequences in the experiment consisted of a set of four sentences, the last of which contained an instance of VPE. The antecedent to the elliptical sentence was contained in the second sentence of the sequence, rather than in the immediately preceding sentence, so simply using information in the most recent sentence or most recent verb phrase would not be an adequate comprehension strategy.

The functionally separated segments of discourse in this experiment took the form of a distinction between dialogue and narrative utterances. This choice allowed the distinction between the utterance types to be created by a single, content-free manipulation: presence or absence of quotation marks. Comparison sequences could be constructed that varied only on whether a particular sentence was surrounded by quotation marks or not, while all other factors were held constant. In the sequences with the functional separation cue, the sentence that intervened between the ellipsis and its antecedent was cast as a narrative remark inserted between two lines of dialogue, as here:

Ellen and her friends were at a barbecue.
 "Someone drank Herb's lemonade," Ellen remarked.
 He had only drunk half of it when it disappeared.
 "Indeed, someone did," Annie agreed.

In the comparison sequences without the functional separations, the last three sentences were all cast as dialogue, as here:

Ellen and her friends were at a barbecue.
 "Someone drank Herb's lemonade," Ellen remarked.
 "He had only drunk half of it when it disappeared."
 "Indeed, someone did," Annie agreed.

In both these cases, "Indeed, someone did" must be understood to mean "Indeed, someone did drink Herb's lemonade" by reference to information from the second sentence. The critical difference between the first sequence above and the second one is the absence of quotation marks around the third sentence in the first sequence, thereby separating it from the dialogue. If utterances that are structurally separated from other parts of the discourse are not considered candidates for use in interpreting the anaphor, then it should be easier for people to locate the needed information in the second sequence than in the first.

The second type of comparison in this experiment was between sequences that contained question-answer pairs and sequences that did not. A typical question-answer sequence was:

Everyone was returning from vacation.
 "Did Greg see Maureen and Marjorie last night?" Helen asked.
 "I think they just got back in town."
 "Yes, he did," Sophia replied.

The comparison sequences without question-answer pairs were simply a context sentence followed by three declarative remarks, similar to the all-dialogue condition of the first comparison:

Everyone was returning from vacation.
 "Greg saw Maureen and Marjorie last night," Helen remarked.
 "I think they just got back in town."
 "Yes, he did," Sophia replied.

In the first sequence above, the second and fourth sentences form a closely related pair, since the question solicits an answer, and the answer can be assumed to refer back to the question. The Declarative sequence lacks this cue. Again the content remains constant between paired declarative and question-answer trials; only the form of the first sentence differs. If people use the relatedness of question-answer pairs, they should be quicker to understand elliptical utterances in sequences of the first sort than the second.

Method

Subjects. Twenty-two Stanford undergraduates received course credit for participating in this experiment. One subject was eliminated from the analyses for excessively long reaction times, and another was eliminated at random to maintain counterbalancing. All subjects in this and the subsequent experiments were native speakers of English.

Stimuli. Each trial consisted of four sentences that formed a short, meaningful sequence, as in the examples given. Trials belonged to one of four conditions, described above. Sixteen Question-Answer trials were constructed with sixteen matched Declarative trials, and sixteen Dialogue-Narrative trials were constructed with sixteen matched All-Dialogue trials. ("Matched" trials were the same sequences of sentences, differing only in presence or absence of the cues.) For all target trials, the fourth sentence contained an instance of VPE, the antecedent to which was located in the second sentence of the trial. The third sentence was a natural-sounding successor to the second, but did not contain a possible antecedent to the elliptical sentence.

All sequences involved a dialogue between two people, as in the examples. For trials with three lines of dialogue (all trials except Dialogue-Narrative), one speaker produced the first two lines of dialogue and a second speaker produced the third in response. For Dialogue-Narrative trials, one speaker produced the first line and another the third, but the sentence between them was not part of the dialogue. Each subject received 32 target trials, 8 from each condition. Mixed with the 32 target trials were 32 filler trials that were constructed to serve two purposes. The first purpose was to prevent subjects from expecting that all trials used the first sentence to interpret a deleted verb phrase in the fourth. To this end, none of the filler trials involved VPE, but contained a variety of other anaphoric devices such as pronominal reference in their fourth sentences. In 22 of the fillers, the anaphor in the last sentence made reference to the information in the third sentence rather than in the second. Combined with target trials, 42 trials involved reference back to the second sentence, and 22 to the third.

The second purpose of the filler trials was to increase the number of trials containing narrative sentences. In the target trials, one quarter of the trials were in the Dialogue-Narrative condition, and the remaining three-fourths of the trials did not contain any narration. The filler trials were therefore weighted toward the Dialogue-Narrative type trials. Combined with target trials, 28 trials contained narrative sentences, and 36 did not.

Filler trials were intermixed with target trials, and the 64 trials were randomly ordered. There were two sets of stimuli,

counterbalanced such that a target trial appearing in one condition in a particular position in one set appeared in its matched condition in the same position in the other set. Each sentence sequence appeared in only one condition in a set; thus, a subject who saw a sequence in Question-Answer form did not also see it in Declarative form.

Procedure. Stimuli were presented on a Hazeltine 1500 CRT. Presentation was controlled by a PDP 11/34 computer running a modified version of the Yale Experimental Psychology System. The four-line stories were presented one line at a time, with display of each line terminated by a button press by the subject. The button press caused the screen to be cleared, reading time for the sentence to be recorded, and the next line in the sequence to be presented. Each sequence of four lines was preceded by the words, "New story. Push button to continue," and each individual line was preceded by a fixation point (a '+') at the left edge of the line.

A true-false question was asked after a random one-quarter of all trials, to encourage subjects to read for comprehension. The question always referred to information in the trial just ended, but it referred equally often to information from the first, second, third, or fourth sentence of the trial. Half the questions were correctly answered "yes," and half, "no."

Subjects were told that the experiment was about how people understand sentences in stories or other texts. They were instructed simply to read each sentence as they normally would for comprehension, and to press the button when they fully understood the sentence. The instructions emphasized that some sentences might be harder than others, and they should be sure they completely understood each sentence before pressing the button. They were also told that a true-false question would be asked after randomly selected stories.

Subjects were given four practice stories before the real trials began. If they gave an incorrect answer to the true-false question in the practice trials, they were asked to read more carefully.

Subjects were given the opportunity to rest after half the trials were over. At the end of the experiment, they were asked whether they had used any particular strategy in reading the sentences or had noticed anything else about the task or their performance.

Results and Discussion

Mean reaction times for the four target conditions are given in

Table 1. These means were calculated eliminating trials with times over 10 seconds or under 100 milliseconds (less than 1% of the data). Results of both primary comparisons were clear-cut. Question-Answer trials were 335 msec faster than the matched Declarative trials, $\text{min } F(1, 32) = 7.72, p < .01$. In contrast, mean reaction times for the Dialogue-Narrative and All-Dialogue conditions were virtually identical; there was an 8 msec difference between the two conditions, which, obviously, was not significant.

 Table 1

Mean RTs, Experiment 1

Question-Answer	Declarative
1.399	1.734
Dialogue-Narrative	All-Dialogue
1.426	1.418

The Question-Answer vs. Declarative comparison suggests that when two utterances are strongly related, as questions and answers are, their relatedness aids in locating the information needed to interpret an anaphor. On the other hand, there is no evidence that readers took advantage of the relatedness cue provided by separation of dialogue and narrative. There are several possible explanations. One is that active expectation of the occurrence of a closely related utterance may be necessary for a benefit to exist. Question-answer pairs are in the class of projecting sequences. A question generates the expectation that a related answer will occur, whereas declarative utterances of either dialogue or narrative do not necessarily generate an expectation about what should follow. Any general expectation that more dialogue will follow a sentence of dialogue may, in fact, be eliminated if the next sentence is of another type. When a reader encounters the structural switch from dialogue to narrative, utterances that were part of the first functional discourse might be discarded from short-term memory as irrelevant to the short-term processing requirements of utterances in the second functional discourse. This is a particular version of an importance-based explanation for retention in short-term memory; it suggests that an utterance is important and will be retained if

it generates an active expectation that a future utterance will depend on it for its interpretation.

The second possibility is that sentences from the first structure can be retained in STM and held for integration with future utterances, but only if it is clear that the first structure is the main thread of the discourse while the second structure is minor. The reason that the dialogue-narrative separation had no effect might then be that a distinction between the main and minor thread was not made in the experimental sequences. This explanation is a more general importance-based interpretation of retention in short-term memory: It suggests that active expectation of closely related utterance pairs is not necessary for retention of particular information, but simply general importance relative to other nearby utterances.

The third possibility is that simple search cues for relatedness normally are what is used, but they were made ineffective by the experimental paradigm. In Experiment 1, the distinction between dialogue and narrative was conveyed simply by presence or absence of quotation marks around one sentence in the sequence. These sentences were written to sound plausible as either part of a dialogue or part of a narrative, and, aside from quotation marks, they were identical in the two conditions. It is possible that subjects were not perceiving the distinction between dialogue and narrative when reading the sentences. They may have tended to read all middle sentences as a continuation of the dialogue, or as narrative, or they may have read some correctly and not others. Since this explanation does not help distinguish between a search-cue and importance-based account, a second experiment was carried out in which the distinction between dialogue and narrative sentences was more obvious.

Experiment 2

While sentences in a dialogue may be very similar to sentences in narrative under some circumstances, there are often features besides quotation marks that distinguish the two. If a participant in a dialogue mentions him or herself, for instance, it will generally be done using "I," as in:

"I hope the trip went well," Mary said.

If that person is later referred to in the narration, then the pronoun will become "he" or "she," as in:

She was wondering if the car held up.

Another distinguishing feature is verb tense. A person reporting an event will often use a present or simple past:

"I want to go out for a few minutes," John said.

or

"I went to the store yesterday," Mary remarked.

while the narrative will often use a past progressive or past perfect:

He was hoping to get Mary to go with him.

or

She had been craving ice cream all day.

Experiment 2 took advantage of these additional cues to unambiguously separate dialogue and narrative in the experimental trials. Sequences in the Dialogue-Narrative condition all contained third sentences with both a person and tense change such that the narrative sentence could only be interpreted as a remark outside the ongoing dialogue rather than as a continuation of the dialogue. Thus, a typical All-dialogue sequence was,

Cathryn was spreading word of a good-bye party for Jim.

"Don't tell Jim about the party," she warned.

"I want it to be a surprise for him."

"I won't," Laura promised.

Its matched Dialogue-Narrative sequence was,

Cathryn was spreading word of a good-bye party for Jim.

"Don't tell Jim about the party," she warned.

She wanted it to be a surprise for him.

"I won't," Laura promised.

If separation of a potential source for an antecedent from intervening parts of the discourse can cue resolution of ellipsis, then, again, Dialogue-Narrative trials should be easier to understand than All-Dialogue trials.

Method

Subjects. Twenty regular and summer school Stanford students participated in this experiment.

Stimuli. Stimuli were sequences of four sentences, as in Experiment 1. The Dialogue-Narrative and All-Dialogue trials of

Experiment 1 were modified for this experiment to make the distinction between dialogue and narrative more salient, as described above. Narrative middle sentences were distinguished from dialogue by absence of quotation marks, as in Experiment 1, but in addition each involved a change of tense and person that clearly set it outside of the dialogue. (Since change of tense sometimes required adding an auxiliary verb, some Dialogue-Narrative middle sentences were longer by this one word than the same sentence in the All-Dialogue sequence. If this small difference had any effect on reading time for the fourth target sentence, however, it would be in the direction opposite to the predicted outcome.)

As in Experiment 1, each stimulus set contained eight Dialogue-Narrative trials and eight All-Dialogue trials. There were again two stimulus sets, one containing the matched All-Dialogue trials to the Dialogue-Narrative trials in the other, and vice-versa. For control purposes, five of the eight Question-Answer and Declarative trials each from Experiment 1 were also included, unchanged. Filler trials were the 32 fillers from Experiment 1. Each stimulus set therefore contained 58 trials, of which the Dialogue-Narrative and All-Dialogue trials were of primary interest.

Procedure. The procedure for this experiment was identical to that in Experiment 1, except that there were fewer total trials per subject.

Results and Discussion

Mean reaction times for the four target conditions are given in Table 2.

Table 2

Mean RTs, Experiment 2

Question-Answer	Declarative
1.328	1.517
Dialogue-Narrative	All-Dialogue
1.387	1.413

The Question-Answer vs. Declarative difference was once again substantial: Question-Answer trials were 189 msec faster than the Declarative trials, $F(1, 18) = 7.98$, $p < .025$. The analysis by items was just short of significant, $F(1, 8) = 5.04$, $p > .05$, but since there were so few items, the statistical power of the test was low. Experiment 1 and an additional comparison of the Question-Answer vs. Declarative sequences in Malt (1982) have provided evidence of the reliability of this comparison. Again, in contrast, the Dialogue-Narrative trials showed an advantage of only 26 msec over the All-Dialogue trials. This difference was not significant.

Thus, there is no evidence in this second experiment that the distinction between dialogue and narrative facilitates finding an antecedent. Three pieces of evidence suggest that this lack of effect is not due to trivial aspects of the experimental context. First, in the two experiments here and an additional one in Malt (1982), the question-answer cue produced large benefits in reading times. Since the stimuli in that comparison differed in construction from those in the Dialogue-Narrative comparison only in the critical variable, and since each subject received all conditions within the same experiment, the apparatus, presentation mode, or use of short 4-line sequences cannot in and of themselves have prevented a benefit from occurring. Second, informal questioning confirmed that the distinction between dialogue sentences and narrative sentences was clear in these stimuli. Subjects reported that they perceived the distinction while reading, and graduate students given a printed version of the stimuli were able to point accurately to the narrative sentences. Finally, since the large effect for question-answer cue and the lack of effect for the dialogue-narrative cue replicated over three experiments, there is no reason to doubt the reliability of the result.

Lack of an effect for the dialogue-narrative comparison must therefore be due to some more fundamental aspect of how these sentences are represented or processed. It appears that separating a sequence of utterances into functionally related parts does not help in interpreting ellipsis; simple relatedness cues do not provide easier access to the necessary information. It is clear, however, that information in utterances two prior to the current one can remain available in memory, since people can understand the question-answer sequences easily. We are thus left with the two possibilities suggested before. One explanation suggests that the utterance containing the antecedent is of insufficient importance in the dialogue-narrative sequences. It is not picked out to remain available in short-term memory as a question is, but rather is lost as new information enters. This explanation suggests, in turn, that if the earlier utterance could

be made more important relative to the one that followed it, it would remain available and comprehension of an elliptical third utterance would be facilitated.

The alternative explanation suggests that perhaps there is something special about the question-answer pair that is not shared by other types of utterance sequences, such as a tendency for a question to generate an active search for an answer in subsequent input. Unless an explicit expectation of dependence exists, connecting an elliptical fragment with a distant antecedent may be difficult.

To some extent, the Dialogue-Narrative vs. All-Dialogue comparison in the second experiment did include what might have been a manipulation of importance of the middle utterance. By changing pronouns from first person to third and verbs from present to past progressive or past perfect, the middle utterance was often changed from what might be considered foreground material to what might be considered background material. If changes in importance were thereby created, then it is unlikely that importance is the critical factor in finding an effect. Hopper and Thompson (1980) point out, though, that unlike in some languages, the background-foreground distinction in English is not explicitly marked by a particle or other device. A given sentence can often play either role and can only be identified as one or the other within a larger context. Since the content of the middle sentence was the same when it was dialogue as when it was narrative, many of the features that distinguish background and foreground sentences did not vary between the two conditions. Thus, although the narrative sentences may have been clearly set apart from the dialogue in this last experiment, they may not have been perceived as of lesser or background importance.

The third experiment attempted to test more directly the notion that a sentence may remain available for subsequent integration, perhaps even as the preferred choice, as long as it is not displaced by a sentence of greater or equal importance. If the question-answer benefit is caused by some factor unique to question-answer pairs or to projecting sequences as a class, then manipulations of general importance should not show any effects. If, however, importance underlies the question-answer effect, then other types of structures that involve distinctions in importance should show reaction time costs and benefits.

This experiment was also designed to test the generality of the conclusion that discourse structure can influence anaphoric processing. One hazard of research using natural language materials is that it is virtually impossible to manipulate a single aspect of the stimuli while perfectly controlling all other possible sources of variation. In changing a sequence from

question-answer to declarative, for instance, the naturalness or typicality of the sequence may also change. One solution to this problem is to show that a variety of different manipulations will produce the same basic result. If it is possible to do so, then the likelihood is low that the result is due to some incidental aspect of the stimuli. Experiment 3 therefore tested two manipulations of structure that differ significantly from each other as well as from the manipulations in Experiments 1 and 2 in order to provide additional support for an effect of discourse structure.

Experiment 3

The two manipulations in this experiment both used non-projecting sequences. Since there are no explicit linguistic devices or functional distinctions in this class to pick out which utterance will contain the antecedent to a later anaphor, no utterance should generate an expectation of specific relatedness or provide a search cue for a guided backward search. If a benefit does occur, it will provide evidence for a general importance interpretation as opposed to an explicit-dependency one and at the same time provide further evidence against a search-cue explanation.

The first comparison in this experiment was between sequences that maintained the same focus throughout versus ones that involved a change of focus. In the Same-focus sequences, the first utterance made a statement or remark, and the middle utterance was, broadly, a remark that elaborated on, justified, or explained the first utterance. Thus, the middle sentence maintained the original focus and was subordinate in importance to the first utterance, as here:

"I liked the Monet exhibit," Deanna remarked.

"It was really very nice."

"I did, too," Richard replied.

In contrast, in the Changed-Focus condition, the middle sentence also plausibly followed the first sentence, but it introduced one or more entities (generally as subject of the sentence) that had not been mentioned in the first sentence. These new entities were part of new primary assertions, rather than simply supporting or evaluating the previous sentence, as here:

"I liked the Monet exhibit," Deanna remarked.

"Renoir is my favorite, though."

"I did, too," Richard replied.

If information is lost from memory by being replaced by information of greater or equal importance, then the first sentence in the Same-Focus sequences should be relatively available for interpreting an elliptical third utterance, but the same sentence in the Change-Focus should be less so.

The second contrast in this experiment was created by either keeping the middle line of the sequence as a separate sentence (as in the previous experiments), or else making it into a subordinate clause of the first. A typical trial of the Subordinate type was,

Carolyn wanted to go skiing over the weekend,
since it was nearing the end of the season.
Peter did, too.

and its matched No-Subordinate trial was,

Carolyn wanted to go skiing over the weekend.
It was nearing the end of the season.
Peter did, too.

The middle line should be treated as more important in sequences in which it is a separate assertion than in those in which it is a subordinate clause. Hence, an importance explanation of retention would predict that elliptical third sentences would be easier to understand in the Subordinate than in the No-Subordinate condition.³

Method

Subjects. Subjects were 16 Stanford undergraduates participating for course credit.

Stimuli. Trials were three lines long, and, as in the previous experiments, the last sentence contained an instance of VPE and the first contained its antecedent. In the Same-Focus vs. Change-Focus comparison, all sentences were dialogue, with the first two sentences spoken by one person, and the third by another. The critical manipulation was in the middle sentence: whether it was an explanation of the first sentence, or a new assertion. Middle sentences of matched trials in the two conditions were always equal in word length, and were constructed such that neither one contained a potential antecedent for the elliptical fragment.

Trials for the Subordinate vs. No-Subordinate comparison were all non-dialogue, since this allowed the use of connectives such as "for" ("...for he had just arrived") that are common in writing but not in conversation. No-Subordinate trials consisted of three separate sentences as usual. Subordinate trials differed

ANAPHORA AND DISCOURSE STRUCTURE

from previous stimuli in that the first line of these trials ended in a comma and was continued in a subordinate clause in the middle line. Seven connectives were used: for, since, because, while, so that, although, and where. Middle lines of the Subordinate condition were identical to those in the Matched No-Subordinate trials, except for lack of capitalization and addition of the connective. The subordinate middle lines were one word longer than No-Subordinate middle lines due to the connective; however, all connectives were short, common words that should require little reading time. Any additional difficulty they did cause in reading the elliptical third sentence would be counter to the predicted result.

Filler trials were 16 dialogue sequences and 16 narrative sequences. Half the filler narrative sequences were three separate sentences, and half were continued from the first line onto the second. The continued ones included a mixture of clause types, some of which were subordinates and some of which were not. The dialogue fillers were taken from previous experiments and were simply three separate lines of dialogue. Filler trials used various forms of anaphora in the third sentence, all referring back to information in the previous sentence.

Procedure. Since the trials for the two comparisons differed quite a bit in form, they were presented in separate blocks. Subordinate and No-Subordinate trials were randomly intermixed with their filler trials and presented in one block, followed by a second block with Same-Focus and Change-Focus trials and their fillers. Subjects were given four practice trials of the Subordinate and No-Subordinate type before the first block of trials and four practice trials of the Same-Focus and Change-Focus type before that block. Otherwise, procedure was identical to the previous experiments. True-false questions were asked after one-quarter of trials.

Results and Discussion

Mean reaction times for the two comparisons are given in Table 3. Reaction times were 314 msec faster for the Same-Focus than for the Change-Focus trials. This difference was significant across subjects, $F(1, 14) = 9.19$, $p < .01$, and across items, $F(1, 14) = 11.79$, $p < .01$, and $\min F'$ was also significant, $\min F'(1, 28) = 5.17$, $p < .05$. Thus an utterance is more likely to remain available if a subsequent utterance is explanatory of the first rather than introducing a new topic.

Subordinate trials showed a benefit of 126 msec over the No-Subordinate trials, and this effect was significant across subjects, $F(1, 14) = 9.49$, $p < .01$. However, the effect was not reliable over items, $F(1, 14) = 1.01$, and $\min F'$ was not

significant.

 Table 3

Mean RTs, Experiment 3

Same-focus	Change-focus
1.316	1.630
Subordinate	No-subordinate
1.284	1.410

It is surprising that the result for Subordinate trials is not stronger, since intuitively elliptical sentences in the Subordinate condition do seem easier to understand than those in the No-Subordinate condition. The one-line-at-a-time presentation mode may be at fault here. Presenting the subordinated clause in isolation after the main part of the sentence may artificially increase its salience and cause it to be read more like a separate assertion. A different presentation mode might therefore yield stronger results for this manipulation.

The difference between Same-Focus and Different-Focus trials does provide further evidence for an effect of discourse structure on understanding anaphora. It also supports the hypothesis that a sentence will remain available until it is replaced by another one of equal importance. How important a subsequent utterance is may be a major factor in determining how long an earlier sentence will be held in short-term memory. Explicit expectancy, as generated by a question, is not essential to retaining a sentence for subsequent integration.

Thus, it appears that certain information is selectively retained in short-term memory, and the information chosen for retention is information that is in some way evaluated as important. The value of such a selection system is presumably that important information is likely to be needed for interpreting subsequent sentences. When an utterance introduces a new assertion or new topic, the information from the older utterance can be put aside because it is no longer the information most likely to be required for subsequent interpretations. If the representation of an utterance is lost from memory when subsequent

input indicates that it will no longer be needed, then other indicators of this need should also influence whether earlier information is retained. Experiment 4 examines this possibility.

Experiment 4

Question-answer pairs were again used in this experiment. When a question is asked, it is logical that it should remain available for subsequent integration. It should remain important until an answer occurs, since interpretation of the answer will usually depend heavily on the question. The first experiments showed that, in fact, questions remained available across intervening sentences to a larger degree than comparable declarative statements did.

In standard question-answer sequences, a question is followed by just one answer. This fact suggests an additional prediction about retention of a question in short-term memory: once an answer is given and interpretation of the answer completed, the question may be dropped from memory. If, however, the question is followed by intervening information that does not provide an answer, then the question should remain available until the answer is given. This prediction differs from a pure importance hypothesis in that in either case the statement following the question may be a response directly relevant to the question, neither being necessarily of lesser importance to the discourse. The two types of response differ simply in whether they do or do not complete the question-answer sequence.

This experiment compared question-answer sequences completed by the sentence directly following the question with sequences that are not completed. Four-sentence sequences were again used in which the first sentence provided a general context, and the remaining three constituted a spoken exchange. In all trials, a question was asked in the first sentence of the dialogue and an answer was given by a different speaker in the last sentence. For half the trials, the middle sentence of the dialogue was cast as an added remark by the asker of the question. Thus, a typical Open sequence might be,

Andrea was anxious to leave the house.
 "Aren't we going to the game?" she asked.
 "It's getting kind of late," she added.
 "Yes, we are," Ron assured her.

In the comparison sequences, the second line of dialogue was spoken by another speaker as an answer to the question asked in the first line of dialogue. Thus a matched trial in the Closed

condition might be,

Andrea was anxious to leave the house.
 "Aren't we going to the game?" she asked.
 "It's getting kind of late," Greg observed.
 "Yes, we are," Ron assured her.

To the extent that people use completion of a question-answer sequence as a cue to drop the question from short-term memory, they should be slower to read elliptical fourth sentences in the Closed than in the Open condition.

Method

Subjects. Twenty-four U.C. Berkeley students served as subjects for course credit.

Stimuli. There was a total of 48 trials, 16 of which were target trials. Half the target trials were in one condition and half were in the other. There were two sets of stimuli, with the trials in one condition for one set in the other condition for the other set. As described above, trials in the Closed condition had a third sentence that was a possible end of the question-answer sequence, while the same sentence in the Open condition functioned only as an explanatory remark.

Middle sentences did not involve elliptical reference back to the question, nor did they contain a potential antecedent to the ellipsis in either condition.

Filler trials were 20 declarative and 12 question-answer sequences. For half of each type of filler, the speaker of the first and second lines of dialogue were the same, with a different speaker for the third, and for the other half, the speaker of the second and third lines were the same with a different speaker for the first. For the question-answer fillers, half the trials of each type had the question in the first line of dialogue and half had it in the second. Since all target trials were question-answer sequences, there were a total of 28 question-answer trials and 20 declaratives. One-quarter of trials were again followed by a true-false question.

Procedure. Procedure was the same as in past experiments, except that subjects were run on an Apple II Plus microcomputer with an Apple video monitor.

Results and Discussion

Mean reaction times for the Open and Closed conditions are given in Table 4. Subjects were 402 msec faster for elliptical fourth

sentences for Open trials than they were for the same sentences in Closed trials. This difference is in the predicted direction and is significant across subjects, $F(1, 22) = 41.24$, $p < .001$, and items, $F(1, 14) = 7.78$, $p < .025$, and $\min F'(1, 19) = 6.54$, $p < .05$.

 Table 4

Mean RTs, Experiment 4

Open	Closed
1.930	2.332

It thus appears that a question remains readily available for interpreting an answer only until an answer is given. Once the question-answer pair is completed, it becomes more difficult to interpret a subsequent utterance as depending on the question. This result provides support for the notion advanced earlier: selection of information for retention in short-term memory is based on likelihood of needing that information for interpreting a subsequent utterance. When that likelihood is low, the information can be dropped from memory.

General Discussion

Several main results have been provided by these experiments. First, and most generally, the structure of an utterance sequence does influence a reader's ability to interpret ellipsis. This fact is shown by differences in comprehension speed for the same elliptical fragment embedded in different types of sequences. Elliptical fragments are easier to understand when the fragment and its antecedent form a question-answer pair than when they form part of a declarative sequence, and when the question has not already received an answer than when it has. They are also easier to understand when information intervening between the fragment and the antecedent explains or evaluates the first utterance, rather than introducing a new topic. They may be easier to understand when the intervening information is a subordinate clause of the utterance containing the antecedent rather than a separate assertion. All these facts demonstrate that the process of resolving ellipsis cannot rely only on

properties of the anaphor or of the individual utterance in which it occurs. Rather, when information in one utterance must be interpreted using information in another one, relations between the utterances will influence ability to arrive at the interpretation. The process of understanding even a single missing verb phrase is tied to the structure of the discourse in which it is embedded. To fully understand how integration at the level of individual words is accomplished, then, it is necessary to consider integration at other levels as well.

The second result is a suggestion of how structure influences the interpretation process. The distinction between dialogue and narrative did not facilitate finding the needed information, while question-answer pairs and sequences maintaining a single focus did. This fact suggests that readers do not use search cues based on separation of a sequence into related parts. Rather, certain information is picked out for retention in short-term memory. This information is then more available for integration than it otherwise would be.

Finally, we also have some idea what this basis for selection is. The system appears to retain information that is likely to be needed for interpreting subsequent utterances. This is illustrated by the types of sequences that produced faster reading times, and in particular by the fourth experiment, which showed that the content of a question becomes less available after an answer is given, when it is no longer critical for interpreting a forthcoming response.

A more general conclusion⁴ is that what information is available for integration is not determined by a strictly first-in-first-out retention system. As mentioned earlier, Kintsch and van Dijk (1978) suggested that hierarchical level of importance to the text as a whole would influence retention of information in short-term storage. Cirilo and Foss (1980) found, however, that higher-level propositions required longer reading time than lower-level propositions. They argued that longer reading times could account for the superior recall of these sentences (Fletcher, 1981), rather than selection for retention in a short-term buffer. The present studies demonstrate, however, that the contents of the buffer are indeed flexible and influenced by importance. While Fletcher's (1981) study used a somewhat artificial cued-recall task to determine what might have been retained longer in short-term memory, the present studies tapped normal comprehension using a task well-established to be sensitive to availability of information in short-term memory; thus, there can be little doubt that short-term availability was being influenced. Readers or listeners seem to generate active expectation about what information will continue to be referred to in a discourse.

One implication of this selective retention concerns the role of focus in interpreting an anaphor. A number of authors have suggested that various types of anaphors are interpreted using information about the focus of a discourse (Grosz, 1977; Sidner, 1978; Yekovitch et al., 1979) or the location of a "discourse pointer" that indicates the current topic (Carpenter & Just, 1977). According to these proposals, a referring expression is generally interpreted as referring to the entity that is the focus of the discourse. This interpretation is achieved through a guided backward search. On encountering an anaphor, the listener or reader performs the basic backward search and preferentially arrives at the focussed entity as the one to use in interpreting the anaphor. A notable feature of these discussions is that they tend to give the focusing process the flavor of a passive marking system. An entity that is marked as focussed has no special role until an anaphor is encountered; the marking is simply used by the backward search in selecting the correct interpretation. In contrast, the present research suggests that these results may, in fact, be come about through the active selection process. An entity identified as the focus of a discourse may be preferentially retained in short-term memory over others that are of only subsidiary importance, making it more available than if it were not the focus.

It is impossible to say which conception is correct on the basis of current results, since both conceptions make the same predictions for standard manipulations of focus. It does seem, though, that discourse focus shares with questions and other retained utterances the critical feature of local importance. This feature is not possessed by other factors that guide anaphora resolution, such as the gender or number of a pronoun, or syntactic constraints on co-reference. These latter cues do not allow any advance predictions about what is likely to be referred back to later; they can only be made use of once the anaphor has occurred. Thus, there may be two classes of factors that influence how readily an anaphor is understood: those that pick out central information to be kept available for integration with future input; and those that operate once an anaphor has been encountered to select among the possibilities in the retained material.

A final implication is for the role of discourse structure in short-term comprehension processes. Early investigations of structural variables in discourse processing concentrated primarily on their influence on long-term memory representations -- how the structure of the discourse influences the structure and content of the long-term representation (e.g. Bartlett, 1932; Mandler & Johnson, 1977; Thorndyke, 1977). More recently, attention has been given to accounting for these long-term representations in terms of processing considerations such as

length of retention in a buffer or reading times (e.g. Cirilo & Foss, 1980; Fletcher, 1981; Kintsch & van Dijk, 1978; Vipond, 1980). Even so, however, there has been relatively little direct concern for the impact of these structural variables on the short-term, sentence-by-sentence process of interpreting the text. The present results illustrate that discourse structure is not something that has its effect only after the basic syntactic and semantic analyses of an utterance (including the resolution of anaphora) have been completed. Instead, it influences the on-line processing of constructing an interpretation. In fact, the local structure examined in this paper probably has much greater influence on this aspect of processing than on ultimate memory for the discourse, since variations at the local level are not necessarily relevant to decisions about long-term storage.

Local discourse structure undoubtedly does more than just facilitate understanding anaphoric reference, of course. When a question is recognized as a question or an assertion as an assertion, information is also conveyed about the speaker or writer's certainty about a proposition, for instance, and about the type of response expected from the listener or reader. It remains to be specified exactly how and when the various types of information are extracted and the decisions about selective retention made. The present results clearly demonstrate, though, that a complete theory of discourse processing will have to include the representation and use of information about utterance types, and about the relations of these utterance types to one another.

Author Notes

This paper is based in part on a dissertation submitted to Stanford University and supported by NIMH grant MH 20021 to Herbert H. Clark. Experiment 4 and preparation of the paper were carried out while the author was supported by a Sloan Foundation fellowship at the University of California, Berkeley. I would like to thank Herbert Clark, Ivan Sag, and Ellen Markman for valuable discussion and guidance, and Herbert Clark, Charles Clifton, and Gregory Murphy for many helpful comments on an earlier draft of this paper.

Footnotes

1. Utterance is used here in the technical sense of a sentence used on a particular occasion for a particular purpose. It refers to both spoken and written forms.

2. The possibilities suggested here do not make any assumption about the sort of representation that is being influenced, other than that it is short in duration and limited in amount. For discussion of the types of representation that may be involved in interpreting anaphors, see Sag and Hankamer (in press; also Hankamer & Sag, 1976.) For experimental evidence bearing on the issue, see Murphy (1983) and Tanenhaus, Carlson, and Seidenberg (in press).

3. This manipulation does, of course, vary whether the antecedent is one or two sentences prior to the elliptical one as it varies subordination. Locating information when a sentence boundary has been crossed may be more difficult than when one has not been, if sentence boundaries are the points at which information is lost from short-term storage. If so, however, this difference may not be so much a confound as part of the same phenomenon. Information can more logically be disposed of at a sentence boundary than at a point within a sentence, since sentences are relatively less dependent upon one another than clauses of the same sentence. An importance-for-future-integration explanation may therefore explain any effect of crossing sentence boundaries, rather than vice-versa. In any case, Clark and Sengul (1979) found that clause rather than sentence boundaries were more relevant to determining how available the referent of a definite noun phrase was, and both Subordinate and No-Subordinate trials here have one clause intervening between the anaphor and its antecedent.

4. Any such conclusions are, of course, most directly applicable to the anaphor used in obtaining the results, VPE. For

arguments for the generalizability of these results, see Malt (1982).

References

- Bartlett, F. C. (1932). Remembering. Cambridge, England: Cambridge University Press.
- Caramazza, A., Grober, E., Garvey, C., and Yates, J. (1977). Comprehension of anaphoric pronouns. Journal of Verbal Learning and Verbal Behavior, 16, 601 - 609.
- Carpenter, P. A., and Just, M. A. (1977). Reading as eyes see it. In M. A. Just and P. A. Carpenter (Eds.), Cognitive Processes in Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Cirilo, R., and Foss, D. (1980). Text structure and reading time for sentences. Journal of Verbal Learning and Verbal Behavior, 19, 96 - 109.
- Clark, H. H. Inferences in comprehension. (1977). In D. LaBerge and S. J. Samuels (Eds.), Basic processes in reading: Perception and Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Clark, H. H. and Sengul, C. J. (1979). In search of referents for nouns and pronouns. Memory and Cognition, 7, 35 - 41.
- Ehrlich, K. (1980). Comprehension of pronouns. Quarterly Journal of Experimental Psychology, 32, 247 - 255.
- Ehrlich, K. (in press). Eye movements in pronoun assignment: A study in sentence integration. In K. Rayner (Ed.), Eye movements in reading: Perceptual and language processes.
- Fletcher, C. (1981). Short-term memory processes in text comprehension. Journal of Verbal Learning and Verbal Behavior, 20, 564 - 674.
- Garrod, S. C. and Sanford, A. J. (1982). The mental representation of discourse in a focussed memory system: Implications for the interpretation of anaphoric noun phrases. Journal of Semantics, 1, 21 - 41.
- Grosz, B. (1977). The representation and use of focus in dialogue understanding. Technical Note No. 151, SRI, Menlo Park.
- Halliday, M., and Hasan, R. (1976). Cohesion in English. London: Longman.

- Hankamer, J. and Sag, I. A. (1976). Deep and surface anaphora. Linguistic Inquiry, 7, 391 - 426.
- Hinds, J. (1979). Organizational patterns in discourse. In T. Givon (Ed.), Syntax and Semantics, Vol. 12: Discourse and Syntax. New York: Academic Press.
- Hobbs, J. R. (1979). Coherence and coreference. Cognitive Science, 3, 67 - 90.
- Hopper, P. J. and Thompson, S. A. (1980). Transitivity in grammar and discourse. Language, 56, 251 - 299.
- Jefferson, G. (1972). Side sequences. In D. Sudnow (Ed.), Studies in Social Interaction. New York: Free Press.
- Kintsch, W. and van Dijk, T. A. (1978). Toward a model of text comprehension and production. Psychological Review, 5, 363 - 394.
- Lesgold, A., Roth, S., and Curtis, M. (1979). Foregrounding effects in discourse comprehension. Journal of Verbal Learning and Verbal Behavior, 18, 291 - 308.
- Longacre, R. E. (1979). The paragraph as a grammatical unit. In T. Givon (Ed.), Syntax and Semantics, Vol. 12: Discourse and Syntax. New York: Academic Press.
- Malt, B. C. (1982). Anaphora and discourse structure. Unpublished dissertation, Stanford University.
- Mandler, J. M. and Johnson, N. S. (1977). Remembrance of things parsed: Story recall and structure. Cognitive Psychology, 9, 111 - 151.
- Marslen-Wilson, W., Levy, E., and Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. Jarvella and W. Klein (Eds.), Speech, Place, and Action: Studies in Deixis and Related Topics. Chichester: J. Wiley & Sons.
- Murphy, G. L. (1982). Understanding anaphora. Unpublished dissertation, Stanford University.
- Murphy, G. L. (in press). Processes in anaphoric understanding. Journal of Verbal Learning and Verbal Behavior.
- Rumelhart, D.E. (1975). Notes on a schema for stories. In D. Bobrow and A. Collins (Eds.), Representation and Understanding: Studies in Cognitive Science. New York:

Academic Press.

- Sag, I. A. (1979). On the nonunity of anaphora. Linguistic Inquiry, 10, 152 - 164.
- Sag, I. A. and Hankamer, J. (1984). Toward a theory of anaphoric processing. Linguistics and Philosophy, 7, 325 - 345.
- Schegloff, E. and Sacks, H. (1973). Opening up closings. Semiotica, 8, 289 - 327.
- Shwartz, S. (1981). The search for pronominal referents. Cognitive Science Technical Report # 10, Yale University.
- Sidner, C. (1978). The use of focus as a tool for disambiguation of definite noun phrases. In D. Waltz (Ed.), Theoretical Issues in Natural Language Processing-2, University of Illinois, Urbana-Champaign.
- Springston, F. (1975). Some cognitive aspects of presupposed coreferential anaphora. Unpublished dissertation, Stanford University.
- Tanenhaus, M. K., Carlson, G. N., and Seidenberg, M. S. (in press). Do listeners compute linguistic representations? In A. Zwicky, L. Karttunen, and D. Dowty (Eds.), Natural Language Parsing: Psycholinguistic, Theoretical, and Computational Perspectives. Cambridge: Cambridge University Press.
- Thorndyke, P. (1977). Cognitive structure in comprehension and memory of narrative discourse. Cognitive Psychology, 9, 77 - 110.
- Townsend, D. J. (1983). Thematic processing in sentences and texts. Cognition, 13, 223 - 261.
- Townsend, D. J. and Bever, T. G. (1978). Interclause relations and clausal processing. Journal of Verbal Learning and Verbal Behavior, 17, 509 - 521.
- Vipond, D. (1980). Micro- and macroprocesses in text comprehension. Journal of Verbal Learning and Verbal Behavior, 19, 276 - 296.
- Yekovitch, F., Walker, C. and Blackman, H. (1979). The role of presupposed and focal information in integrating sentences. Journal of Verbal Learning and Verbal Behavior, 18, 535 - 548.