ANAPHORA RESOLUTION: SHORT-TERM MEMORY AND FOCUSING

RAYMONDE GUINDON

Microelectronics and Computer Technology Corporation
(MCC)
9430 Research Blvd.
Austin, Texas 78759.

ABSTRACT

Anaphora resolution is the process of determining the referent of anaphors, such as definite noun phrases and pronouns, in a discourse. Computational linguists, in modeling the process of anaphora resolution, have proposed the notion of focusing. Focusing is the process, engaged in by a reader, of selecting a subset of the discourse items and making them highly available for further computations. This paper provides a cognitive basis for anaphora resolution and focusing. Human memory is divided into a short-term, an operating, and a long-term memory. Short-term memory can only contain a small number of meaning units and its retrieval time is fast. Short-term memory is divided into a cache and a buffer. The cache contains a subset of meaning units expressed in the previous sentences and the buffer holds a representation of the incoming sentence. Focusing is realized in the cache that contains a subset of the most topical units and a subset of the most recent units in the text. The information stored in the cache is used to integrate the incoming sentence with the preceding discourse. Pronouns should be used to refer to units in focus. Operating memory contains a very large number of units but its retrieval time is slow. It contains the previous text units that are not in the cache. It comprises the text units not in focus. Definite noun phrases should be used to refer to units not in focus. Two empirical studies are described that demonstrate the cognitive basis for focusing, the use of definite noun phrases to refer to antecedents not in focus, and the use of pronouns to refer to antecedents in focus.

INTRODUCTION

The goal of this research is to show the relation between the psychological work on anaphora resolution based on the notion of a limited short-term or working memory and the computational linguistics work based on the notion of focusing.

This rapprochement is important for the following reasons:
1) From a theoretical viewpoint, cognitive evidence increases the validity of the computational notion of focus.
2) Focusing corresponds to one of the reader's comprehension processes and it needs to be incorporated in the model of the user in language understanding systems to adequately resolve ambiguities in the user's utterances and to handle language generation.

FOCUSING IN COMPUTATIONAL LINGUISTICS

According to Grosz (1977), who was interested in the resolution of definite noun phrases, focusing is the process, engaged in by participants in a discourse, of highlighting a subset of their shared reality. Grosz, Joshi, and Weinstein (1983) distinguish between two levels of focus, global focus and centering. Global focusing is a major factor in maintaining global coherence and in the interpretation of definite noun phrases. Centering is a major factor in maintaining local coherence and in the interpretation of pronouns.

Grosz, Joshi, and Weinstein further define the notion of centering. Each sentence has two types of centers whose purpose is to integrate the sentence to the discourse. The backward-looking center links the current sentence to the preceding discourse. The set of forward-looking centers provides the set of entities to which further anaphors may refer. The backward-looking center corresponds, roughly, to Sidner's focus and the forward-looking centers to Sidner's potential foci.

One principle derived by Grosz, Joshi, and Weinstein is the following: if the backward-looking center of the current utterance is the same as the backward-looking center of the previous utterance, a pronoun should be used. In other words, if there are no topic shifts, continue to refer to the same entity by using a pronoun.

However, violations of this principle have been presented in Grosz (1977) and noted in Grosz, Joshi, and Weinstein (1983). They have shown that pronouns are sometimes used to refer to entities mentioned many sentences back, even though the backward-looking center of intervening sentences has been changed by topic shifts.

Sidner (1979, 1983) has proposed the notion of focus in the context of interpreting anaphors, especially pronouns. In Sidner's theory, an anaphor neither refers to another word nor co-refers to another word, but rather co-specifies a cognitive element in the reader's mind. Moreover, a theory of anaphora resolution must predict the pattern of reader's correct and incorrect choices of co-specifiers and the failures to understand. This view makes explicit the consideration of the reader's mental model and inferential capabilities.

A sketch of Sidner's focusing process follows. First, an initial focus is selected on the basis of syntactic features and thematic roles indicating topicality in the first sentence. Other elements introduced in the sentence are stored as potential foci for later sentences. When an anaphoric expression is encountered, this focus is tested as a co-specifier for the anaphor. It has to satisfy syntactic restrictions on co-references (Lasnik, 1976), semantic selectional restrictions (Katz and Fodor, 1963), and pragmatic plausibility constraints expressed in the remainder of the sentence. If the focus fails as a co-specifier for the anaphor, the potential foci are tried in turn. At the same time, the new elements introduced in the sentence are stored as potential foci for later sentences. Third, the focus is updated to the selected co-specifier for the anaphor. If the focus has changed, a topic shift has occurred. The second and third steps are cyclically applied after each sentence.

The advantage of using a focus mechanism is that it prioritizes and restricts the search for a co-specifier, and as a consequence, reduces the computational costs associated with inferential processing when testing the applicability of the co-specifier to the anaphor.

COGNITIVE STUDIES OF ANAPHORA RESOLUTION

A few representative empirical studies of anaphora resolution are described below. All the experimental paradigms used share the following assumptions:
1) human memory is functionally or structurally divided into at least two types of memories, a short-term memory with small storage capacity but very fast retrieval time and a long-term memory with very large storage capacity but slow retrieval time;
2) a topic shift transfers the units currently in short-term memory to long-term memory;
3) an anaphor transfers its referent from long-term memory to short-term memory (i.e. reinstates its referent), if it was not already in short-term memory.

The first assumption is crucial. Other things being equal, computations involving retrieval from short-term memory will be faster than those involving retrieval from long-term memory. Turning to the second assumption, topic shifts have been found to be induced with a varity of linguistic devices. One of the devices is the introduction of intervening sentences between the referent and its anaphor. The intervening sentences are unrelated to the referent but related to the overall text. Another device is the specification of a temporal or spatial parameter that is outside the normal range of a situation. When describing a dinner, the phrase "Five hours later," signals that the topic of conversation is no longer the dinner. Another device is the use of an anaphor, frequently a definite noun phrase, to refer to an antecedent that is not currently the topic of conversation but is in the "background". Finally, there is the use of key phrases to signal a diversion in the flow of discourse, such as "Let's turn to.", as documented in Reichman (1978, 1984).

The general pattern for the material used in these experiments is the following. At the beginning of the text appears a sentence containing a referent (e.g. biologist). For example, "The mission included a biologist". Then, if the referent should not be in focus, the next sentence or sentences indicate a topic shift as described above (e.g.

unrelated intervening sentences). If the referent should be in focus, no devices for topic shifts are used. The following sentence then contains an anaphor (e.g. scientist, he) to the focused or non-focused referent (e.g. biologist). For example, "The scientist collected samples from the cultures". Another example is shown in Table 1 of this paper.

Carpenter and Just (1977) used eye tracking with other converging techniques to study anaphora resolution. With eye tracking, one can monitor very precisely the trajectory of the eyes, with their forward and regressive movements, and the duration of eye fixations on small segments of the text. The assumption behind using this technique is that eye movements are closely related to higher level cognitive activities such as comprehension. Therefore, one can expect longer fixation durations on text segments requiring additional processing to be comprehended and one can expect the eye movement pattern to mirror the selective pickup of important information in the text.

They performed a series of experiments testing the effect of recency of a referent on the time course of anaphora resolution. Indirectly, they tested the effect of recency on the availability of an item in short-term memory. They presented texts where the number of sentences between the referent and the anaphor was varied from zero to three. The subjects read each sentence and, after the sentence, had to decide whether it was consistent or inconsistent with the previous sentences. The consistency judgment times and the eye fixations were recorded. The consistency judgment task, used as converging evidence with the eye movement technique, is believed to induce the subjects to integrate each new sentence and should parallel the difficulty of anaphora resolution. The overall reading time of the anaphoric sentence was measured using the eye tracking technique. Each of these tasks should be faster if the referent was in short-term memory than if the referent was in long-term memory.

Response times for the consistency judgments and reading times of the anaphoric sentences increased as the number of intervening sentences increased. The sharpest difference appeared between zero and one intervening sentence. Gaze durations within the anaphoric sentence were shorter when there were no intervening sentences than in the other conditions.

These results show not only that anaphora resolution is easier when the referent is nearer the anaphor but also that one intervening sentence may be sufficient to produce a topic shift.

Clark and Sengul (1979) used the sentence reading time technique to study anaphora resolution. In this technique, subjects control the onset and offset of the presentation of a sentence by pressing a button. The subjects are instructed to press the button to see a new sentence as soon as they have understood the current sentence. The assumption behind this technique is that additional processing required for comprehension will increase sentence reading time.

Clark and Sengul (1979) measured the reading time of a sentence containing an anaphor. They distinguished between two models of the effect of recency of a referent on the speed of anaphora resolution. In the first model, called the "continuity model", entities mentioned in the discourse are searched backward from the last one. One should expect monotonically increasing reading time as the searched entity is farther back. In the second model, called the "discontinuity model", entities mentioned in the current or last sentence are kept in short-term memory and accessed first. All the entities that are further back are more likely to be in long-term memory (and not in short-term memory) and accessed second.

Subjects read short paragraphs where a referent could be separated from the anaphor by zero to two intervening sentences. The reading time of the sentence containing the anaphor was fast when the referent was in the immediately preceding sentence but equally slow when it was two or three sentences before. This finding supports the discontinuity model. Entities in the last processing cycle are more likely to be kept in short-term memory than entities in previously processed cycles. Once a text entity is not in short-term, the number of intervening sentences does not affect the speed of anaphora resolution.

Lesgold, Roth, and Curtis (1979), who related the linguistic notion of foregrounding (Chafe, 1972) to the psychological notion of short-term memory, performed a series of experiments similar to those of Clark and Sengul (1979), using more varied ways to produce topic shifts, and replicated the above findings.

McKoon and Ratcliff (1980) used an activation procedure based on Chang (1980). A description of the basic paradigm and its underlying logic follows. When one reads a text, only a small part of the text information is stored in short-term memory and most of the information is stored in long-term memory. This is due to the very small storage capacity of short-term memory (7 ±2 chunks; Miller, 1956). Given that retrieval time in short-term memory is much faster than retrieval time in long-term memory, it will take longer to remember something from the text if the memory is stored in long-term memory than in short-term memory.

In their study, subjects read a paragraph sentence by sentence. Immediately after the last sentence, the subjects were presented with a single word and the subjects had to remember whether the word had appeared previously in the text or not (an old-new recognition). If the tested word was still in short-term memory, the old-new recognition time should be faster than if it was in long-term memory.

To test this hypothesis, the paragraphs were constructed in the following manner. The referent (e.g. burglar) was separated from the anaphor by either zero or two intervening sentences. The anaphor appeared in the last sentence of the paragraph. The last sentence was presented in one of three versions: 1) the subject of the sentence was a repetition (i.e. burglar) of the referent in the first sentence (anaphoric-identical); 2) the subject was the name of the category (e.g. criminal) in which the referent belonged (anaphoric- category); 3) the subject was a noun (e.g. cat) unrelated to the referent (non-anaphoric). During the experimental trials, the "referent" (i.e. burglar) was presented immediately after the last sentence for an old-new recognition.

Assuming that an anaphor activates its referent by making it available in short-term memory, one can expect significantly faster old-new recognition times for "burglar" in the anaphoric-category condition than in the non-anaphoric condition. This prediction was observed.

Surprisingly, the number of intervening sentences did not have an effect. This suggests that the two intervening sentences did not remove the referent from short-term memory (i.e. "backgrounds" the referent). It is probably not the case. Rather, it is likely that by testing the referent at the end of the clause, as opposed to when the anaphor is encountered, the referent had time to be reinstated in short-term memory and be highly available. This is an important point. The activation procedure was <u>not</u> on-line since the old-new recognition occured at the <u>end</u> of the sentence as opposed to <u>while</u> the sentence was read and the anaphor encountered.

Another initially surprising effect was that the old-new recognition times for the referents were slower in the zero intervening sentences when the anaphor was a repetition of the referent itself than when the anaphor was the category name. This last result suggests that it is not appropriate to use a definite noun phrase, especially a repetition of the referent, to refer to a antecedent in short-term memory.

As explained previously, intervening sentences are not the only devices that transfer text units from short-term to long-term memory. Stereotypical situations have spatial and temporal parameters with legal ranges of values. If one specifies a spatial or temporal value outside these ranges, a scenario-shift occurs. For example, Anderson (in Sanford and Garrod, 1981) constructed texts about stereotypical situations such as going to a restaurant. In one sentence of the text, there was a reference to a character related to the script, say a waiter. At the beginning of the next sentence, there was a mention of a temporal or spatial parameter, such as "One hour later" or "Five hours later". In the first case the parameter is within the range defining the script, in the second case it is not. The rest of the sentence contained an anaphor to the previously mentioned character, the waiter. Measuring the reading time of the anaphoric sentence, Anderson showed longer reading time when the spatial or temporal parameter was outside the range of the script than inside. This suggests that the referent was transfered from short-term to long-term memory by the scenario-shift and it took longer to retrieve the referent during anaphora resolution.

The results from all these experiments support the notion that an anaphor activates its referent by making it highly available in short-term memory and that topic shifts transfer units from short-term memory to long-term memory. However, none of these studies, except some eye movement studies, provide data on <u>when</u> anaphora resolution occurs during the reading of a sentence and when it occurs in relation to the

lexical, syntactic, semantic, and pragmatic analyses.

COGNITIVE BASIS FOR FOCUSING

A sketch of a cognitive model of anaphora resolution is offered here. It has been heavily influenced by the short-term/long-term memory model of Kintsch and van Dijk (1978) and especially its leading edge strategy.

Structure of human memory

Analogically, human memory can be conceptualized as a three level structure similar to the memory of most mini and main frame computers. It consists of a small, very fast memory called short-term memory (STM); a relatively larger main or operating memory (OM); and a vast store of general world knowledge called long-term memory (LTM).

The total STM is only large enough to contain 7±2 chunks of information at any one time (Simon, 1974; Miller, 1956). The resources for STM are dynamically allocated to one of two uses. First, part of the STM is used to store the incoming sentence or clause. This is a temporary storage of the sentence or clause before further processing and is called the STM buffer. The second part of STM is called the STM cache. It is used to hold over, from one sentence or clause to the next, the information necessary to provide local and global coherence. It contains a subset of the previous text items that are topical and a subset of those that are recent. Retrieval times from short-term memory are very fast.

Conceptually, operating memory is the subset of the world knowledge in long-term memory which is deemed relevant to the processing of the current part of the text. It also contains the growing memory structure corresponding to the text read so far. It contains the less topical and less recent information from the text. Retrieval times are much longer than for short-term memory.

Processes

The time course of anaphora resolution is greatly determined by the current content of short-term memory and of operating memory. Moroever, pronouns and definite noun phrases are resolved using different strategies.

Cache management. During the input of a sentence into the buffer and the concurrent integration of the sentence into the cache, a subset of the semantic units held in the STM is selected to be held over in the cache for the next cycle. Following Kintsch and van Dijk (1978), the cache management strategy selects a subset T of the most topical items and a subset R of the most recent items to be held over in the cache. The selection strategy aims at maximizing the probability that an anaphor in the next sentence will refer to a semantic unit held in the cache. Cache management is applied after each sentence or clause.

Anaphora resolution - Integration. Pronouns and definite noun phrases are resolved using different strategies. We will describe four cases:

1. The anaphor is a definite noun phrase and the referent is not in focus, that is, it is in operating memory.

2. The anaphor is a definite noun phrase and the referent is in focus, that is, it is in the cache.

3. The anaphor is a pronoun and the referent is in the cache (in focus).

4. The anaphor is a pronoun and the referent is in operating memory (not in focus).

It is hypothesized that the explicitness of an anaphor is a signal, used by the reader, which denotes whether the referent is in the cache or in operating memory.

If the anaphor is a definite noun phrase, operating memory is searched immediately. If the referent is in operating memory it is then reinstated into the cache. A topic shift has occured.

If the anaphor is a definite noun phrase and the referent is in focus (i.e. in the cache), anaphora resolution will be hindered. The reader searches operating memory while the referent is in short-term memory. Correspondingly, this violates a rule of cooperative communication: use a definite noun phrase to refer to an antecedent not in focus. The definite noun phrase signals a topic shift, while in fact, the same entity is being talked about.

If the anaphor is a pronoun, the cache is searched for a plausible referent. If found, anaphora resolution is completed. Because cache management is based on topicality and recency, pronouns can refer to the main topic of the text even when the main topic has not been mentioned directly for many sentences. Unless there is a global topic shift, the main topic in the cache remains unchanged throughout the text.

If the anaphor is a pronoun but no referent is found in the cache, it is then necessary to search operating memory. If a referent is found in operating memory, it is reinstated into the cache. A topic shift has occured. Using a pronoun to refer to information in operating memory is detrimental to anaphora resolution. The reader first searches the cache, then the operating memory, and then has to reinstate the referent into the cache.

COMPARISONS

A clear relation exists between the notion of focusing proposed in computational linguistics and the model of human memory and discourse processing proposed in cognitive psychology.

The cache is used to store the items in focus. Given the small number of items stored in the cache, a sketchy anaphor such as a pronoun is sufficient to retrieve the referent. The cache management strategy in human memory is aimed at maximizing the probability that the cache contains the information relevant to the next cycle of computation. The cache, by containing topical and recent items, allows to maintain global and local coherence.

Operating memory is used to store items that are not in focus. Because the set of items is large, an informative description of the item to be searched for is needed. Definite noun phrases are used to indicate to the reader that the item is not in focus, thus in operating memory. Other things being equal, it will take more time to retrieve an item from operating memory than from the cache. The referent will need to be reinstated into the cache. This will produce a topic shift. The reinstated referent is then highly available and can be referred to by using a pronoun.

TWO ON-LINE STUDIES OF ANAPHORA RESOLUTION

The presented studies test the notion that focus is cognitively realized in the reader's limited short-term memory. They also test Grosz, Joshi, and Weinstein's claim that definite noun phrases, and not pronouns, should be used to refer to items no longer in focus and that pronouns, and not definite noun phrases, should be used to refer to items in focus. Moreover, if one assumes that the content of short-term memory is dynamically updated on the basis of recency and topicality, one can explain why pronouns can be used to refer to recent items and also to topical non-recent items.

A new technique, called on-line activation, was developed specifically to provide the empirical data for these studies. The on-line activation technique can be compared to "closely" tracing the execution of a program.

In the on-line activation technique, passages are presented using rapid serial visual presentation (RSVP), one word at a time. In addition to reading each text, the participants were also given the task to recognize whether some specially marked words, presented surreptitiously within the text, had appeared before in the text or not. Some of these special words were presented before in the text and others were not. We will call these specially marked words test words. This task is called an old-new recognition task.

The passages contained anaphors referring to antecedents which were either in focus or not in focus. An antecedent was removed from focus by introducing a topic shift, with the restriction that the antecedent was not the main topic of the discourse. An example text is presented in table 1. Note that only one of the alternative sentences 5a, 5b, or 5c was presented for each text to the participants of the study.

In each text, one of the test words was the referent of the anaphor. At some point before or after the anaphor was presented on the CRT, its referent was presented for old-new recognition and recognition times and errors were collected. The delay between the onset of the anaphor and the onset of the test word is called the stimulus onset asynchrony (SOA). The anaphor is acting as a prime, which should activate the referent. The old-new recognition time for the referent test word indicates

whether the referent is in the cache or in operating memory.

TABLE 1

EXAMPLE OF TEXTS WITH ANTECEDENTS IN FOCUS AND NOT IN FOCUS

Antecedent: thermometer
Anaphor: instrument

Antecedent in Focus

1- The assistant was preparing solutions for a chemistry experiment.
2- The experiment would take at least four hours.
3- There would then be a ten hour wait for the reaction to complete.
4- He measured the temperature of a solution using a thermometer.
5a- The thin instrument was not giving the expected reading.
5b- A broken instrument was not giving the expected reading.
5c- The computer terminal was not giving the expected reading.

Antecedent not in Focus

1- The assistant was preparing solutions for a chemistry experiment.
2- He measured the temperature of a solution using a thermometer.
3- The experiment would take at least four hours.
4- There would then be a ten hour wait for the reaction to complete.
5a- The thin instrument was not giving the expected reading.
5b- A broken instrument was not giving the expected reading.
5c- The computer terminal was not giving the expected reading.

In addition, there were three types of primes, as shown in sentences 5a, 5b, and 5c in Table 1. The prime could be either semantically related and referential (S+R+) as in 5a, semantically related and not referential (S+R-) as in 5b, or semantically unrelated and not referential (S-R-) as in 5c. In the S+R+ condition, the prime is the anaphor. The two conditions S+R- and S-R- were control conditions to separate the effect of semantic priming, due to semantic association between the anaphor and the referent, on the old-new recognition for referents.

A schema of the procedure is shown in Table 2. The words surrounded by stars are the test words.

TABLE 2

SCHEMA OF THE PROCEDURE

SOAs	Before	350 msec	1250 msec
Time			
T1	The	The	The
T2	thin	thin	thin
T3	*thermometer*	instrument	instrument
T4	instrument	*thermometer*	was
T5	was	was	not
T6	not	not	giving
T7	giving	giving	*thermometer*

The predictions were:

1. If a referent is not in focus, due to a topic shift, the occurence of the anaphor should reinstate the referent into the cache, leading to faster old-new recognition times. In terms of the experimental conditions, there should be a decrease in old-new recognition time at the 350 and 1250 msec conditions in the S+R+ condition (i.e. after the anaphor), but not in the S+R- and S-R- conditions, which are not anaphoric.

2. The use of a definite noun phrase to refer to an antecedent in the cache (i.e. in focus) should be detrimental to anaphora resolution. It should slow down the recognition of the referent as old or new. In terms of the experimental conditions, if the referent is in focus, the old-new recognition times in the 350 and 1250 msec SOA conditions should be slower than in the before SOA condition.

Method

Participants There were 36 participants in this study.

Materials There were 36 experimental texts. They contained as a referent an instance of a class (e.g. thermometer) to be used later as a test word, and as an anaphor the class name (e.g. instrument). In this study, the anaphor was a definite noun phrase. An example of the material was presented in Table 1. There were three priming conditions, S+R+, S+R-, and S-R-,

exemplified respectively by sentences 5a, 5b, and 5c.

During the presentation of each text, two or three test words were presented, one experimental and one or two fillers. The filler words were presented at semi-random locations in the text. In the entire experiment there was an equal number of old and new test words.

Procedure The experiment was computer-controlled using real-time routines on the VAX/VMS 11/780 of the Computer Laboratory for Instruction in Psychological Research at the University of Colorado. Each participant sat in front of a CRT screen with a keyboard which had a "yes" button on the right, for old test words, and a "no" button on the left, for new test words. The texts were presented using RSVP, with each word presented in the center of the screen for 300 msec. The participants were asked to recognize whether the test words were old or new, as fast as possible but without making mistakes.

Design There were 36 experimental texts and 18 experimental conditions. The first manipulation was the focusing of the referent: in focus or not in focus. The second manipulation was the SOA: immediately before the prime, 350 msec after, 1250 msec after. The third manipulation was priming: S+R+, S+R-, S-R-. The design was completely within-subject, with two texts randomly assigned to each experimental condition using two randomly sampled 18 by 18 Latin Squares. Each participant was randomly assigned to a row of the Latin Squares.

Results and Discusion

The predicted interaction of focusing and priming is shown in Figure 1: the prime in the S+R+ condition (i.e. the anaphor) reinstates the referent into the caohe, focusing it, while the referent is not reinstated in the non-referential conditions, $F(2,70) = 3.6$, $p < 0.04$, MSe = 213421 by subjects and $F(2,70) = 2.5$, $p < 0.09$, MSe = 277568 by items. A priori comparisons show that the difference between the recognition times in the two focus conditions in the S+R+ condition is much smaller than in the other two priming conditions, S+R- and S-R-, which do not differ between themselves, $t(35) = 2.6$, $p < 0.01$, MSe = 87 by subjects, and $t(35) = 2.14$, $p < 0.02$, MSe = 114 by items. These results support the notions that items in focus are more accessible than items not in focus and that focus is realized into the cache. They also support the notion that an anaphor reinstates a referent not in focus and does so by transferring the referent to the cache.

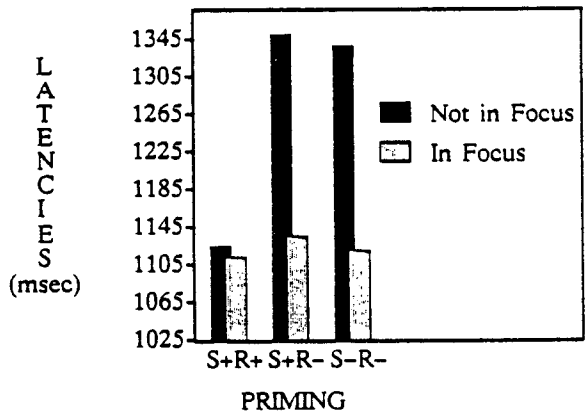

FIGURE 1. Recognition latencies at each focus and priming condition.

An a priori comparison demonstrates that using a definite noun phrase to refer to an item in focus hinders anaphora resolution. What seems to happen is a surprize effect caused by the violation of a linguistic usage relating the form of the anaphor to the focus status of its referent. The recognition time for the referent, in the focus condition, was longer at the 350 msec and 1250 msec SOAs than in the before SOA, $t(35) = -4.1$, $p < 0.001$, MSe = 24 by subjects, and $t(35) = -2.9$, $p < 0.006$, MSe = 31 by items. This is shown in Figure 2.

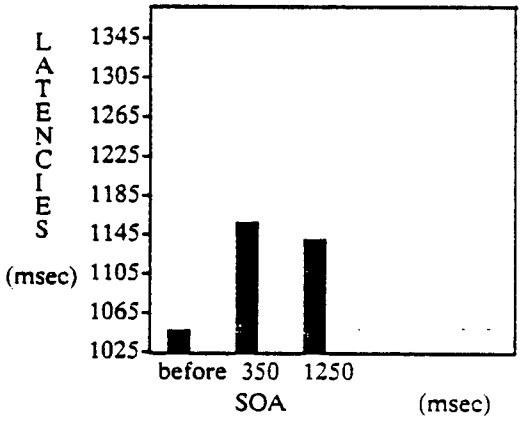

FIGURE 2. Recognition latencies at each SOA for a referent in focus.

In another study (Guindon, 1982), using the same on-line activation technique, the activation of an antecedent by a pronoun was traced. In this study, it was found that referring to an antecedent not in focus by using a pronoun was detrimental to anaphora resolution. The delay between reading the anaphor and reinstating the antecedent was as long as 2400 msec. The activation of an antecedent not in focus by a pronoun takes a long time because the reader is induced: 1) to search the cache unsuccesfully; 2) to search operating memory with a "sketchy" pronoun; 3) to reinstate the referent into the cache. Activation was immediate for the antecedents in focus. As opposed to the previous study where referring to a focused referent using a definite noun phrase hindered anaphora resolution, no such effect was observed when using a pronoun. This is expected since pronouns signal that the referent is in the cache.

SUMMARY

The notion of focusing and the notion that the form of the anaphor signals whether the referent is in focus or not have cognitive support. Items in focus are items in the cache which is dynamically updated to contain the T most topical and the R most recent items in the text. Because the cache contains few items, pronouns should be used to refer to items in focus. Other things being equal, anaphora resolution will be easier if the antecedent is in focus, because the retrieval times from the cache are much faster than those from the operating memory. Antecedents not in focus are in operating memory. Items not in focus are in operating memory. A definite noun phrase, because it is more descriptive than a pronoun, should be used to retrieve the antecedent from the large set of items in operating memory. However, because the retrieval time is slow in operating memory, anaphora resolution is more difficult for items that are not in focus. The reinstatement of an antecedent into the cache effects a topic shift.

The on-line activation technique was developed specifically to provide empirical data on the notion of focus. The advantage of this technique over conventional memory experiments is that one can test precisely the temporal properties of various analyses and processes occurring during sentence and text comprehension. This technique can be used to distinguish between different models of anaphora resolution when these models are not easily distinguished on the basis of discourse or dialogue analysis.

REFERENCES

Carpenter, P.A. & Just, M.A. Integrative processes in comprehension. In D. LaBerge & S.J. Samuels (Eds.), Basic Processes in Reading. Hillsdale, N.J.: Erlbaum, 1977.

Chafe, W. Discourse structure and human knowledge. In J.B. Carroll & R.O. Freedle (Eds.), Language Comprehension and the Acquisition of Knowledge. Washington: Winston, 1972.

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

Clark, H.H., & Sengul, C.J. In search of referents for nouns and pronouns. Memory and Cognition, 1979, 7, 35 - 41.

van Dijk, T.A. & Kintsch, W. Strategies of Discourse Comprehension. New York: Academic Press, 1983.

Grosz, B.J. The representation and use of focus in dialogue understanding. Technical Note 151, Artificial Intelligence Center, SRI, 1977.

Grosz, B.J., Joshi, A.K., & Weinstein, S. Providing a unified account of definite noun phrases in discourse. Technical Note 292, Artificial Intelligence Center, SRI, 1983.

Guindon, R. On-line processing of pronouns: Evidence for reinstatement searches. Unpublished manuscript. University of Colorado, Boulder, 1982.

Guindon, R. The effect of recency and topicality of referents on the time course of anaphora resolution. Doctoral Dissertation. University of Colorado, Boulder, 1984.

Just, M.A. & Carpenter, P.A. A theory of reading: From eye fixations to comprehension. Psychological Review, 1980, 87, 329 - 354.

Katz, J.J. & Fodor, J.A. The structure of a semantic theory. Language, 1963, 39, 170 - 210.

Kintsch, W. & van Dijk, T.A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363 - 394.

Lasnik, H. Remarks on co-reference, Linguistic Analysis, 1976, 2, 1-22.

Lesgold, A.M., Roth, S.F., & Curtis, M.E. Foregrounding effects in discourse comprehension. Journal of Verbal Learning and Verbal Comprehension, 1979, 18, 281 - 308.

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

Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81 - 97.

Reichman, R. Conversational coherency, Cognitive Science, 1978, 2, 283-327.

Reichman, R. Extended person-machine interface. Artificial Intelligence, 1984, 22, 157 - 218.

Sanford, A.J. & Garrod, S.C. Understanding Written Language. New York: Wiley, 1981.

Sidner, C. Towards a computational theory of definite anaphora comprehension in English discourse. Technical report 537. MIT Artificial Intelligence Laboratory, Cambridge MA, 1979.

Sidner, C. Focusing in the comprehension of definite anaphora. In M. Brady and R. C. Berwick (Eds.), Computational Models of Discourse, Cambridge: MIT Press, 1983.

Simon, H.A. How big is a chunk? Science, 1974, 183, 482 - 488.

ACKNOWLEDGMENT

This research was performed as part of the author's doctoral dissertation while at the University of Colorado. She is extremely grateful for the help of her dissertation committee, Walter Kintsch, Peter Polson, Alice Healy, Richard Olson, Andrzej Ehrenfeucht. Burton Wagner has provided many insightful comments on this paper. MCC is kindly thanked for the technical support provided while composing this paper.