1997

The role of multiple antecedents in the time course of anaphor resolution.

Robert A. Mason

University of Massachusetts Amherst

THE ROLE OF MULTIPLE ANTECEDENTS IN THE TIME COURSE OF ANAPHOR
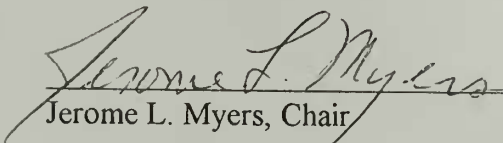
RESOLUTION.

A Thesis Presented

by

ROBERT A. MASON JR.

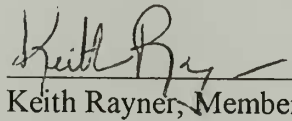# THE ROLE OF MULTIPLE ANTECEDENTS IN THE TIME COURSE OF ANAPHOR RESOLUTION.
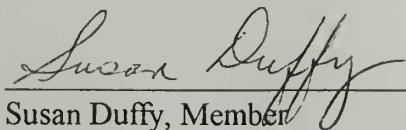
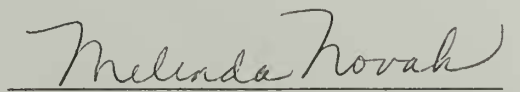A Thesis Presented

by

ROBERT A. MASON JR.

Approved as to style and content by:

_____
Jerome L. Myers, Chair

_____
Keith Rayner, Member

_____
Gary Marcus, Member

_____
Susan Duffy, Member

_____
Melinda Novak, Department Head
Department of Psychology

ACKNOWLEDGEMENTS

I would first like to thank the members of my committee for the time and assistance they gave me in completing this thesis. I would like to thank Gary Marcus for helping me to not lose sight of the broader picture. Thanks to Susan Duffy for helping me understand the workings of the Duffy and Rayner model. Thanks to Keith Rayner for his assistance in all aspects of the eye-tracking experiment and for the shock of a fifteen minute meeting. Also, I would like to thank all the members of the cognitive community for offering ideas when I presented this work at evening seminar.

In addition I would like to thank all the graduate students and post-docs who have been both friends and co-workers. A special thank you to the people who have helped me with my eye-tracking data and listening to my ideas at every stage of my masters: Jason Albrecht, Kathy Binder, and Martin Fischer. Thanks to my fellow workers of the 10pm. to 2am shift, Erik Reichle, Sascha Englebrecht and the night janitors. Thanks to all my friends who have helped me to fill up my time outside of Tobin: Doug, Lori, Helene and Brian. Thanks also to Kelly Schmitt and Dan McCall for always being around. Thanks to Loukia Loukopolous for being the best of friends and someone who I could always argue with. And, though they are not around anymore, my honorary big sisters, Celia Klin, for showing me the way, and Jill Lohmeier, who I could always talk to about anything and everything.

A special thank you to my family for supporting me in all ways while I'm here, though they might not know what I do. Just as helpful to me during this time has been my second family, my housemates, Erik Coats, Lucy Robin and Dan Robin. They helped me to see cognitive

psychology from different viewpoints and have a family away from Pittsburgh. Dan especially has helped me to see that even at this point in life when you are working on a difficult task, there is still time to create a great friendship.

Most importantly, I would like to thank Jerry Myers, my advisor. Jerry has been the best advisor I could have hoped for. He has spent endless hours of his time to help me in all aspects of this research. Through long research meetings he has helped me to begin developing into a scientist, by thinking like one. He has helped me start to write like a professional and use a semicolon; I can't imagine what he thought after reading through my early drafts. He has helped me in the interpretation of the data by explaining aspects of statistical analysis, developing programs for 'crunching' the numbers, and debugging all my programs after I spent hours staring at them not working. Jerry has been both an advisor and a friend. He has pushed me when necessary and at the same time he has been willing to discuss sports especially UMASS basketball. For all these things he deserves my thanks.

ABSTRACT

THE ROLE OF MULTIPLE ANTECEDENTS IN THE TIME COURSE OF ANAPHOR RESOLUTION

FEBRUARY 1997

ROBERT A. MASON JR., B.S., UNIVERSITY OF PITTSBURGH

M.S., UNIVERSITY OF MASSACHUSETTS, AMHERST

Directed by: Professor Jerome L. Myers

In two experiments, participants read passages containing one or two candidates for an anaphoric reference that differed in their distance from the reference. In Experiment 1 eye movements were recorded. The presence of a second potential antecedent caused readers to look longer at the anaphoric reference and regress to a disambiguating region. A recognition probe experiment showed facilitation of the appropriate antecedent. More distant antecedent slowed reading times and resulted in slower recognition judgements.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION AND LITERATURE REVIEW

The goal of research in discourse processing is to arrive at a model that describes, explains and predicts what occurs during comprehension. Any model that attempts to do this must account for the reader's ability to comprehend pronominal and categorical references. Most models of text comprehension describe the process of resolving such anaphoric references in terms of activation; it is assumed that the antecedent is selected that acquires the most activation after the reader encounters the referent (e.g., Corbett & Chang, 1983; Dell, McKoon & Ratcliff, 1983; Kintsch, 1988; McKoon & Ratcliff, 1980; O'Brien, Duffy and Myers, 1986). However, we still lack a complete understanding of this process of activation and selection, and of how it is affected by such factors as distance between the anaphor and its antecedent, the presence of other potential antecedents (or "distractors"), and the semantic relation between the anaphor and both the antecedent and any distractors. In the experiments reported here, we investigate the roles of distance and multiple antecedents in the anaphor resolution process.

Anaphors are actually a type of inference in which some person or concept that has been previously mentioned in the text base is referred to again. Rayner and Pollatsek (1989) state that "in all cases when a word is an anaphor, the reader has to decide which previous concept is being referenced, and in most cases decide whether a previous concept is indeed being referenced" (p.269). Consider the following examples:

(1)   Bill was walking down the street.

        He tripped over the curb.

(2)   Bill was walking down the street with Mrs. Jones.

        He tripped over the curb.

In both cases a connection has to be made between the referent in the second sentence and its antecedent in the first sentence.

In (1) above, connecting the referent he to the antecedent Bill is relatively easy. Even when a second referent is added, the connection is made relatively quickly provided that gender information is unambiguous, as in (2). In a probe experiment, Corbett and Chang (1983) found that, at the end of the sentence, the pronoun's antecedent was more active than the nonantecedent. Gernsbacher (1989) showed that even after immediately encountering the pronoun, the nonantecedent was suppressed when gender information was unambiguous.

This selection process, even though it results in suppression immediately after the pronoun is encountered, still takes time. In order to demonstrate this, Vonk (1984) compared gender-unambiguous with gender-ambiguous sentences. In the gender-ambiguous condition, she used sentences in which the implicit causality of the verb helped to determine pronoun assignment, as in the following example:

(3)   Bill handed John some tickets to a concert,

        but he took the tickets back immediately.

Although a connection was established between the anaphor and its antecedent in both types of sentences, Vonk's results were consistent with the conclusion that assignment was

delayed in the gender-ambiguous condition. She found that, in this condition, fixation time on the pronoun was shorter than when the pronoun was unambiguous with respect to gender. This result may seem counter-intuitive but it makes sense if we consider the following explanation. When gender information is unambiguous, the reader gazes at the pronoun long enough to make the connection between the anaphor and its antecedent. In contrast, when gender information is ambiguous the reader is unable to select the correct antecedent and quickly continues reading, possibly expecting some disambiguating information in the next clause. This is also consistent with the result that Vonk found on the regions following the pronoun. There, she found that in the gender ambiguous condition, fixation times were longer on the verb following the pronoun than when gender was unambiguous. Also, second pass fixation times (i.e., fixations before and after a regression was made) were longer on the sentence as a whole when gender was ambiguous than when gender was unambiguous.

Gernsbacher (1989) used recognition probes to test for suppression effects on materials that were similar to those that Vonk used. Suppression effects are demonstrated if the recognition response to a probe word after an anaphoric reference to that word is slower than a base-line recognition response to the same word. Gernsbacher found that the distractor was suppressed in the gender ambiguous case, but suppression was delayed until the end of the sentence. This result, together with Vonk's (1984), suggests that anaphor resolution may be a two-stage process in which selection between potential antecedents occurs after an initial access of several possible antecedents. In cases in which gender is unambiguous, selection can be made in the first stage, as seen in the slow-down

3

on the pronoun itself But when gender information is ambiguous, selection cannot be made until after a second stage, as evidenced by the slow-down on the region following the pronoun.

Because pronouns cannot be used to refer to distant antecedents, most studies of pronoun resolution have used single sentences or sentence pairs. In contrast, noun anaphors, categorical nouns which refer to some concept previously mentioned, allow for anaphoric references to be made after several sentences. Consider the following sentences:

(4)   A bus went down the street.

        The vehicle smashed into a tree.

(5)   The bus just missed hitting the broken-down car.

          Soon after, the stalled vehicle started.

Although the two sentences are consecutive in both (4) and (5), it is possible to insert several sentences between them and still make the reference clear. In these two examples, the anaphor is no longer a pronoun and is now either the <u>vehicle</u> or <u>the stalled vehicle.</u> When readers encounter the word <u>vehicle,</u> they then have to connect it to its antecedent in the previous sentence, that is, the bus, in (4), or the broken-down car, in (5). This results in a slowdown in reading times when the anaphoric inference is made (e.g., Clark and Sengul, 1979; Garrod and Sanford, 1977; O'Brien, Duffy and Myers, 1986), together with faster responses to antecedent probes (e.g., Dell, McKoon and Ratcliff, 1983; O'Brien et al., 1986).

In contrast to research on pronoun anaphors, there have been few studies of the time course of noun anaphor resolution. Two exceptions are studies by Duffy and Rayner

4

(1990), and Corbett (1984), who examined reading times for the anaphor and post-anaphor region. Both studies obtained evidence consistent with the two-stage process model mentioned previously. Their findings, and the Duffy and Rayner process model, will be discussed in more detail later.

The variables in the present experiments, were distance between the anaphor and its antecedent, presence of a distractor, and position of the antecedent relative to that distractor. All of these variables have been shown to affect the antecedent retrieval process. Let us first examine the distance effect.

Several investigators (e.g. Clark & Sengul, 1979; Ehrlich & Rayner, 1983) have found that anaphors were resolved more quickly as the distance between the anaphor and the antecedent was decreased. Ehrlich and Rayner found that there was no difference between near and intermediate conditions but that the fixations on the two words following the pronoun were longer in the far condition than in the near and intermediate conditions. Clifton and Ferreira (1987) investigated this distance effect and found that reading times on an anaphor were equal for the different distances as long as the antecedent was in focus. Recall that Ehrlich and Rayner only found a distance effect in their far condition; in this condition, the antecedent was separated from the anaphor by a sentence that introduced a new topic. It appears that Ehrlich and Rayner's distance effect was actually a shift-intopic effect.

The effects of distance are somewhat more complicated with longer passages. In several experiments using noun anaphors, O'Brien and his colleagues have found a distance effect (O'Brien, 1987; O'Brien and Myers, 1987; O'Brien, Plewes, and Albrecht,

1990). The materials in these studies consisted of long passages in which the antecedent appeared either early or late in the passage; neither the early nor late antecedent was still in working memory when the anaphor appeared. Sentences containing an anaphoric reference to the late antecedent were read faster than those containing an anaphoric reference to the more distant, early antecedent (O'Brien, 1987; O'Brien, Plewes, & Albrecht, 1990). Such a result has been taken as evidence for a backward search from the anaphor (O'Brien, 1987). However, for several reasons it is not clear that physical distance is the source of this effect. First, the early antecedent was not only more distant from the anaphor; it was also always followed by another potential antecedent. Thus physical distance and relative position were perfectly correlated. Second, other factors, related to the substance of the text, appear to play a more important role. O'Brien (1987) found that importance, as measured by a rating task, was correlated with search time. For example, when a distant early antecedent was rated as more important than a corresponding late antecedent, the early antecedent was reinstated more quickly than the late antecedent. Consistent with this, O'Brien & Myers (1987), using the same passages as O'Brien (1987), found that the number of causal connections to an antecedent was the only significant predictor of recall times in their two experiments. Third, when O'Brien & Myers varied the amount of material between early and late antecedent, thus increasing the distance to the early antecedent, they obtained no effect.

In summary, it can be assumed that when the antecedent was in focus in both the near and far conditions, there was no distance effect (Clifton and Ferreira, 1987; Duffy and Rayner, 1990). Similarly, evidence from long passages (O'Brien, 1987; O'Brien and

Myers, 1987) shows that there was no distance effect when both the near and the far antecedents were outside of the working memory span and other textual factors were held constant. It was only when the far antecedent was out of working memory and the near antecedent was still in working memory that a distance effect occurred (Ehrlich and Rayner, 1983; Duffy and Rayner, 1990). The distance effect obtained by Duffy and Rayner will be discussed in detail later.

Corbett (1984) has shown that the presence of a distractor slows the anaphor resolution process. He measured word-by-word reading times to investigate how the anaphor resolution process was affected by the presence of categorically consistent non-antecedents, or distractors. In the first of two experiments, Corbett measured the overall reading times for the anaphor region, which consisted of an adjective-noun phrase (e.g. frozen vegetable) and the three words which followed it. These overall reading times were longer for those passages in which a distractor was present. Breaking down the overall reading times, Corbett found that reading times were equal for the adjective in both the distractor-present and distractor-absent passages. This was also true for the last word of the reference string in both conditions. Thus, the slow-down in reading times due to the presence of a distractor first occurred on the category noun and continued over the next two words.

In Corbett's second experiment, he varied the presence or absence of a second candidate antecedent (a distractor) and typicality of the antecedent and distractor. For example, when the reference was "frozen vegetable," the two possible antecedents were "frozen asparagus" or "fresh corn." Using the same procedure as in Experiment 1, Corbett

once again found a distractor interference effect; however, there was an interaction with typicality. When the antecedent was a highly typical exemplar and the distractor was atypical, Corbett did not obtain an interference effect. In contrast, when the antecedent was atypical and the distractor was highly typical, there was a large interference effect. In short, interference was a function of the relative activation of the antecedent and the distractor. Assuming that a more typical exemplar will have a higher level of activation, there should be a greater interference effect when it is a distractor and less when it is the antecedent, which is what Corbett found. Although there was no main effect of typicality, such an effect may have been hidden by the long reading times that resulted from the word-by-word technique. The long reading times are caused by a lack of peripheral information that is available during normal reading. This might serve to eliminate any effects that are short in duration and may be difficult to detect when reading times are longer and more variable.

Using longer passages, O'Brien, Plewes, and Albrecht (1990) also investigated the effects of distractors upon anaphor resolution. They determined that the distractors slowed reading times for the sentence containing the anaphor only when the distractor was from the same semantic category. Combined with Corbett's results, this indicates that distractors do interfere with the anaphor resolution process. However, this slow-down only occurs when the distractors are from the same semantic category. Furthermore, the slow-down will be on a continuum; typical distractors will cause an even greater slow-down than atypical ones.

Finally, when there are distractors, it is possible that their relative position with respect to the antecedent may affect the resolution process. Such an effect would be consistent with the backward search hypothesis put forth by O'Brien (1987). However, evidence on this point is unclear. Although Corbett (1984) did not obtain evidence that there was an influence of position of distractor on reading times, the lack of an effect may be due to the use of a small number of items. Corbett only used eight passages, four of which did not have a distractor, which leaves four passages, two with antecedent before distractor and two with antecedent after a distractor. As mentioned earlier, the O'Brien et al. (1990) experiments covaried relative position with distance, which in turn was correlated with causal relatedness. However, the results are suggestive of a positional effect. It is possible that when there are distractor effects, they interact with other variables such as relative position, semantic category, and typicality. The lack of a controlled study of positional variables is part of the motivation for Experiment 1, in which presence of a distractor and position of an antecedent were manipulated.

In summary, the variables discussed above -- distance of antecedent, presence of a distractor, position of the antecedent relative to the distractor, and typicality of the antecedent and distractor -- need to be considered in anaphor resolution. Distance between the anaphor and its antecedent does not affect the resolution process except if there is a change of topic between the antecedent and anaphor, or if both the anaphor and antecedent are out of working memory and other textual factors (i.e. elaboration, causal connection) are held constant. Presence of a distractor slows anaphor resolution if the distractor is from the same semantic category as the antecedent. It also has a greater effect

when the distractor is a typical exemplar of the anaphor's category than when the antecedent is an atypical exemplar. When there are distractors present, their position with respect to the antecedent may also become important. The distractor may cause less interference when it appears before the antecedent than when it appears after the antecedent (O'Brien et al., 1990) but as noted previously, the evidence is inconclusive. Typicality of the antecedent and the distractor may also affect the resolution process; more typical exemplars should result in faster processing times and will probably interact with the previously listed variables. The variables reviewed have different effects upon times on the anaphor region and on the region following the anaphor (Corbett, 1984; Duffy and Rayner, 1990). This finding suggests that anaphor resolution involves more than a single stage. Yet we still lack an understanding of when a distractor affects the anaphor resolution process and how it does so. We need to determine if having multiple candidates for an anaphor affects some kind of access of the candidates or selection of accessed candidates or both. We next consider a proposed two stage model for anaphor resolution. This two-stage model will provide a framework in which we can discuss the effects of multiple antecedents.

CHAPTER II

DUFFY AND RAYNER'S TWO-STAGE MODEL

Duffy and Rayner used noun anaphors that appeared either one line after (near) or four lines after (far) an antecedent. Using anaphors that were category names and antecedents that were exemplars of those categories, they varied typicality of the exemplar and distance between the anaphor and the antecedent. Gaze duration on the target noun was the same in all conditions except when the antecedent was near and was a highly typical exemplar; in this condition, gaze durations were shorter than in the other three conditions. There was also a typicality effect on the post-target region: readers had shorter gaze durations when the antecedents were typical exemplars of the anaphor's category than when they were atypical exemplars. Based on this pattern of reading times, Duffy and Rayner hypothesized that anaphor resolution occurs in two stages; (1) an identification stage, in which lexical access is completed and one or more possible antecedents are identified; and (2) a verification stage, in which the actual antecedent is verified as the intended referent.

According to Duffy and Rayner, the first stage is reflected in the gaze duration times on the target noun (i.e. the anaphor) itself Recall that they found shorter gaze durations when the antecedent was close and typical than in any other condition. They argued that this effect did not reflect facilitation of lexical access for the target word, but rather a facilitation of the identification stage of the anaphor resolution process. This priming occurred as a result of a recent encounter with a word that was semantically

related to the target anaphor. Priming of lexical access was ruled out as a possibility because it would involve priming across sentence boundaries which conflicts with evidence that priming does not occur across clausal boundaries (Carroll and Slowiaczek 1986), nor across sentence boundaries (Garrod and Sanford 1977). It is only when an integrative process is necessary, as in anaphor resolution, that priming occurs across boundaries. Thus, Duffy and Rayner concluded that their priming effect was not a priming of lexical access of the anaphor but priming of an early stage of the resolution process.

The verification stage was hypothesized to be reflected in the gaze durations on the post-target region. The typicality effect observed in the post-target region provided evidence that anaphor resolution was not completed in the time the eyes spent on the target noun but continued during the fixations on the following words. One possible explanation is that this second stage was simply a spill-over of processing from the first stage; however, for this to be true, each stage should be affected by the same factors. This is not what occurred. The gaze durations on the target noun suggest that anaphor resolution was affected by an interaction of distance and typicality, but in the post-target region only a typicality main effect was found. The fact that the times in the two regions were influenced by different factors indicated that two different stages were involved.

The general framework underlying this research involves the following assumptions derived from the Duffy and Rayner Model:

(1)   Anaphor resolution is a two-stage process,

(2)   The first stage is an identification stage which is measured by the gaze duration on the target anaphor,

(3)   High typical/near antecedents prime the identification stage,

(4)   The second stage is a verification stage (a selection stage may also be necessary if more than one antecedent is accessed) which is measured by the gaze duration on the region following the target anaphor.

Within this framework, the nature of the processes involved in each stage still needs to be specified. Experiments 1 and 2 are designed to further our understanding of these processes and how the presence of distractors, and distance to the antecedent affect them.

The primary purpose of Experiment 1 was to address the following question: When there is a distractor present, at what point does it affect the resolution process? The design of Experiment 1 was a replication of Corbett's Experiment 1 (1984) using an eyetracker as opposed to word-by-word reading and using more passages for additional power. This allowed us to address the previous question based on the pattern of times over versions of text. In Experiment 2, readers were probed for recognition of the antecedent or distractor after reading an anaphoric reference statement or no anaphoric reference statement. This enabled us to be sure resolution had occurred, even in the presence of a distractor, and provided a converging measure of access of antecedents.

CHAPTER III

EXPERIMENT 1

A. Introduction

Experiment 1 involved the manipulation of the variables of distance, either a near or far antecedent, and distractor, either absent or present. When there was no distractor present, and because only high typical items were used, we expected to replicate the Duffy and Rayner (1990) distance effects. Near antecedents should be accessed faster than far antecedents in the first stage, and there should be no difference due to distance in the second stage. Given the results obtained by Corbett (1984) we expected to find an effect due to presence of a distractor. The purpose of Experiment 1 was to determine at what point in the process of reading this distractor effect occurs.

The distractor effect should emerge either on the target noun alone, on the post-target region, or at both locations. First, assume that the distractor affects gaze durations on the target noun alone. Such an effect would be consistent with a model in which it takes longer to identify more than one antecedent. This would suggest an effect in the identification stage of the Duffy and Rayner model.

An effect in the target noun region would also be consistent with models in which an identification process, together with a selection process to select among the identified antecedents, both occur during the gaze duration on a target noun. The distractor's presence could affect the identification and/or the selection stage.

It is also possible that a distractor effect will appear in the post-target region. There are two types of process models that would be consistent with such an effect. The first is based on the assumption that all candidates are identified while the eyes are on the target noun, and a selection stage occurs after the eyes have made a saccade to the post-target region; there would be an increase in time spent on the post-target region due to the presence of a distractor. Alternatively, suppose only one antecedent is identified during the time the eyes spend on the target noun. This also may result in a distractor effect on the post-target region. Assume that the first candidate antecedent to acquire activation above a threshold level is the distractor. The distractor might be identified first if it were the near candidate antecedent. In the subsequent verification stage, verification will fail and cause longer gaze durations in the post-target region and possible regressions due to the need to re-engage the identification stage. This failure is a result of the mismatch between the adjective that describes the antecedent and the adjective that describes the anaphor. Such an effect in the verification stage is similar to what Sanford and his colleagues (Sanford, Garrod, Lucas and Henderson, 1983) termed "inappropriate bonding." Thus, under this unique identification model, the distractor effect would be expected to occur only in the far antecedent condition and only in the post-target region.

Recall that Corbett (1984) found that when a distractor was present, longer reading times first appeared on the target noun and then continued into the post-target region. Consistent with this would be the finding that presence of a distractor increases gaze durations on both the target noun and the post-target region. Such a result would suggest a model in which identification is assumed to take longer when there is more than

one possible antecedent and a selection stage occurs while the eyes are fixated in the post-target region.

In summary, Experiment 1 will address the question: when a distractor is present, at what point does it affect the anaphor resolution process? Once we can determine when multiple candidate antecedents have their effect, we can then consider possible models of the resolution process. Furthermore, Experiment 1 should provide a better understanding of the effects of distance to an antecedent, presence of a distractor, and position of a distractor. The results of manipulating these variables will allow us to develop a better understanding of the anaphor resolution process and perhaps to further develop the Duffy and Rayner model.

B. Method

1. Participants

Forty-one members of the University of Massachusetts community were paid or received experimental credit towards psychology classes for participation in the study. Of those, only twenty-eight were used in the experiment. The others were discarded either due to an inability to successfully track where their eyes were fixated in the experiment, or to a large number of tracking losses that produced data that were impossible to analyze.

2. Apparatus

Eye movements were recorded by a Stanford Research Institute Dual Purkinje Eyetracker which has a resolution of 10' of arc. The eyetracker was interfaced with an American Computer Innovations 486 computer which ran the experiment. Viewing was

binocular, with eye location recorded from the right eye. The position of the participant's eye was sampled every millisecond by the computer and averaged over four consecutive samples. The averaged horizontal and vertical positions of the eye were compared with those of the previous sample to determine whether the eye was fixated or moving.

Passages were presented on a View Sonic 17G monitor, with up to 72 character spaces per line. During the experiment, the participant was seated 62 cm. from the monitor, where four characters equal 1 degree of visual angle. The characters were presented in lower case except when upper case was called for (at the beginning of sentences and proper names). Luminance on the monitor was adjusted to a comfortable brightness level for the participant then held constant. The room was dark except for an indirect light source that enabled the experimenter to keep notes during the experiment.

3. Materials

Twenty-four experimental passages were constructed. Each paragraph was approximately eight to nine lines long, with approximately 70 characters per line. The target noun always appeared in the center of a line with the post-target region occupying the remainder of that line.

Each participant read 24 experimental paragraphs, six in each of the four conditions created by crossing distance to the antecedent (near or far) and distractor (present or absent). Each participant was randomly assigned to one of four sets of materials. Each participant read six passages in each condition, and across groups each passage appeared once in each condition. Each set of 24 passages appeared in a larger set of 53 passages. Approximately one quarter of the passages were followed by a

comprehension question. The antecedent appeared on line four, the far position, or on line six, the near position. The line in which the antecedent did not appear contained either a distractor or a noun phrase that cannot be a possible antecedent. Thus the four possible conditions are:

Antecedent	Distractor	Line 4	Line 6
Near	Present	antecedent	distractor
Near	Absent	antecedent	no distractor
Far	Present	distractor	antecedent
Far	Absent	no distractor	antecedent

The target noun and the post-target region appeared in line seven. The target noun phrase (the determiner "the," an adjective and the category name) always appeared in the center of the seventh line. The post-target region consisted of those words in line seven that followed the target noun phrase. This region never included the end of a sentence. An example paragraph is presented in Table 1.

Table 1

Sample Passage

Far Antecedent / Distractor Present
   John and Lucy were walking through the park on a sunny
   day in the middle of June. There were two boys throwing
   a football back and forth under the trees. One of the boys
   was standing under a giant oak and as he threw the ball, Lucy
   yelled, "hey, throw it here!" The other boy, turned from his
   position under the shady elm and threw the ball to Lucy.
   Lucy passed it back through the giant tree to the first boy.
   The boys waved as John and Lucy continued on.

Near Antecedent / Distractor Present
   John and Lucy were walking through the park on a sunny
   day in the middle of June. There were two boys throwing
   a football back and forth under the trees. One of the boys
   was standing under a giant oak and as he threw the ball, Lucy
   yelled, "hey, throw it here!" The other boy, turned from his
   position under the shady elm and threw the ball to Lucy.
   Lucy passed it back through the shady tree to the first boy.
   The boys waved as John and Lucy continued on.

Far Antecedent / Distractor Absent
   John and Lucy were walking through the park on a sunny
   day in the middle of June. There were two boys throwing
   a football back and forth under the trees. One of the boys
   was standing under a giant oak and as he threw the ball, Lucy
   yelled, "hey, throw it here!" The other boy, turned from his
   position near the end of the park and threw the ball to Lucy.
   Lucy passed it back through the giant tree to the first boy.
   The boys waved as John and Lucy continued on.

Near Antecedent / Distractor Absent
   John and Lucy were walking through the park on a sunny
   day in the middle of June. There were two boys throwing
   a football back and forth under the trees. One of the boys
   was standing near the end of the park and as he threw the ball,
   Lucy yelled, "hey, throw it here!" The other boy, turned from
   his position under the shady elm and threw the ball to Lucy.
   Lucy passed it back through the shady tree to the first boy.
   The boys waved as John and Lucy continued on.

4. Procedure

When a participant arrived for the experiment, a bite bar was prepared to eliminate head movements, and the eye-tracker was calibrated. The initial calibration procedure took approximately 5 min. Prior to reading each passage calibration of the eye tracking system was checked to ensure that accurate records were obtained. Each participant read 3 practice passages followed by the set of 24 experimental passages and 26 filler passages. Participants were told that they would be reading a series of paragraphs displayed on a CRT screen. They were told to read for comprehension so that they were able to answer an occasional comprehension question.

At the beginning of each trial five boxes appeared on the screen, and the participant was instructed to look at the left-most box when the experimenter said "ready." Once the experimenter had determined that the participant was fixating the box, the entire passage was presented on the screen to begin the trial. When the participant was finished reading the passage, he or she was instructed to press a button that would end the trial. Participants were asked comprehension questions on approximately 25% of the passages, to which they responded by pressing either a "yes" button or a "no" button. No individual items were excluded from analysis on the basis of answers to comprehension questions.

C. Results

Processing time was measured for three regions: the determiner "the" plus the adjective, the categorical noun anaphor, and the three or four words that followed the anaphoric reference. The first region will be referred to as the disambiguating region, the second as the target region, and the third as the post-target region. The post-target region

did not include the end of a sentence nor did it extend past the end of a line onto the next line. The measures that are reported for these regions are gaze duration, total time, and regressions into a region. Gaze duration is the sum of all fixations on a word before the eyes leave that word. This measure does not include any regressions to the specified region from other parts of the text. Total time consists of the gaze duration for a specific region plus any time the eyes spend in that region after a regression from any other part of the text. Regressions are the number of times the eyes return to a region after first going past it; they are reported in probability of making a regression.

For each measure in the three regions, a 2 (distractor present v. distractor absent) x 2 (far antecedent v. near antecedent) analysis of variance (**ANOVA**) was performed, with tests against both subject (F_1) and item (F_2) variability. In all analyses reported, an alpha level of .05 was used.

1. Disambiguating Region

Because the disambiguating region varied in length, each gaze duration was divided by the number of characters in the word to yield a measure of gaze duration in milliseconds per character. The means for gaze durations, total time and probability of a regression are presented in Table 2.

Table 2

Mean Processing Times and Probability of Regressions for the Disambiguating Region as a Function of Distance of Antecedent and Presence of a Distractor in Experiment 1.

Condition	Gaze Duration[a]		Total Time[b]		Regressions[c]
	Mean	SD	Mean	SD	Mean
Far-Present	32.09	7.81	362	116	13.7
Near-Present	30.37	7.74	344	95	11.9
Far-Absent	31.28	6.75	320	97	6.5
Near-Absent	30.17	5.45	318	71	8.3

Note. [a]Processing times are reported in milliseconds per character. [b]Processing times are reported in milliseconds per word. [c]Regressions are reported as probability.

The gaze durations on the disambiguating region were not significantly affected by the conditions. Processing time was the same regardless of presence of a distractor or distance from the antecedent. An ANOVA indicated no significant effects.

Total time (gaze duration + duration of fixations after a regression) spent in the region did vary across conditions. Readers spent more time in the disambiguating region if a distractor was present, $F_1(1,24) = 5.746$, MSE $= 5,604$; $F_2(1,20) = 6.174$, MSE $= 3,449$. The distance x distractor interaction did not approach significance ($F < 1$), nor did the effect of distance ($p > .20$).

Because the distractor increased total time but did not affect gaze duration, regressions were analyzed to determine if readers looked back more often when there was a distractor present, or if they took more time on those occasions when a regression was made. Readers were more likely to return to the disambiguating region with a distractor present than when there was only one candidate antecedent, $F_1(1,24) = 6.271$, MSE = 0.013; $F_2(1,20) = 3.582$, MSE = 0.014, $p < .07$.

2. Target Region

As with the disambiguating region, gaze durations are presented in milliseconds per character because the target words varied in length. The means of the measures reported for the target region are presented in Table 3.

Table 3

Mean Processing Times and Probability of Regressions for the Target Region as a

Function of Distance of Antecedent and Presence of a Distractor in Experiment 1.

Condition	Gaze Duration[a]		Total Time[b]		Regressions[c]
	Mean	SD	Mean	SD	Mean
Far-Present	50.42	17.38	311	78	4.8
Near-Present	43.59	8.39	298	77	4.2
Far-Absent	45.35	11.93	307	107	4.8
Near-Absent	41.81	9.99	286	56	3.0

Note. [a]Processing times are reported in milliseconds per character. [b]Processing times are

reported in milliseconds per word. [c]Regressions are reported as probability.

Gaze durations were longer when the antecedent was far than when the

antecedent was near; this effect was significant in the subjects analysis, $F_1(1,24) = 9.616$,

MSE = 78, and nearly so in the items analysis, $F_2(1,20) = 3.865$, MSE = 5,373, p = .063.

Readers also spent more time on the target word when a distractor was present than when

one wasn't; this effect was significant when tested against subject variability, $F_1(1,24) =$

4.399, MSE = 74, but not when tested against item variability. The interaction of

distractor x distance was not significant (p>.2).

Neither total time spent in the target region, nor probability of regressions to the target region differed across conditions.

3. Post-Target Region

There were no significant effects in the analysis of the post-target region data. This was true of all three measures: gaze duration, total time, and probability of regression as presented in Table 4. There was one difference on a measure not reported for the previous regions.[1] First fixations on the post-target region were longer when the antecedent was far than when it was near, this effect was significant when tested against subject variability, $F_1(1,24) = 5.371$, MSE = 1,045, but not when tested against item variability. First fixations are meaningless when a region includes more than one word because they do not reflect processing on the region as a whole. However, it has been suggested that they are evidence of a spill-over of processing from a previous region (Balota, Pollatsek, and Rayner, 1985; Rayner and Duffy, 1986; O'Brien, Shank, Myers, and Rayner, 1988). Thus, this result suggests that there is some spill-over of processing when the antecedent is more distant.

Table 4

Mean Processing Times and Probability of Regressions for the Post-Target Region as a

Function of Distance of Antecedent and Presence of a Distractor in Experiment 1.

Condition	Gaze Duration[a]		Total Time[b]		Regressions[c]
	Mean	SD	Mean	SD	Mean
Far-Present	29.18	6.69	539	137	12.5
Near-Present	30.67	9.24	520	154	7.7
Far-Absent	30.12	8.31	503	126	8.3
Near-Absent	30.61	6.51	522	129	10.7

Note. [a]Processing times are reported in milliseconds per character. [b]Processing times are

reported in milliseconds per word. [c]Regressions are reported as probability.

D. Discussion

There are several results from Experiment 1 that contribute to our understanding

of the anaphor resolution process. Distractors disrupted the reading process as evidenced

by three findings. First, there were longer gaze durations in the target region when a

distractor was present. Second, there was a higher probability of regressions to the region

that disambiguated the anaphoric reference when a distractor was present. Third, there

was an increase in total time spent in the disambiguating region with a distractor present.

That the total time effect is due to the regression effect can be deduced from the fact that

there was no difference in gaze durations between distractor present versus distractor

absent conditions in the disambiguating region. It was only after the eyes had left the region and came back later that a difference due to the presence of a distractor emerged. Manipulation of the distance variable also affected the anaphor resolution process. As the distance between antecedent and anaphor increased, gaze durations on the target noun increased accompanied by a spillover into the post-target region.

In order to provide a framework for discussing these effects and a hypothesis about selection of antecedents, let us return to the Duffy and Rayner (1990) model. According to their model, gaze durations on an anaphor reflect lexical access of the anaphor and an initial identification stage of the anaphor resolution process. This identification stage identifies one or more candidate antecedents for the anaphor. The eyes then move past the anaphoric noun and a verification stage is entered. During the verification stage, a candidate antecedent is verified as the intended reference for the anaphor. Duffy and Rayner (1990) did not use passages that contained multiple candidate antecedents; thus there was no need to include a selection stage in their model. This selection process is distinct from either the identification process or the verification process. What remains is to propose a model of anaphor resolution that specifies when this selection stage occurs and how the identification and selection stages function.

With evidence that the distractor effect appeared on the target region alone, we can reject several types of anaphor resolution models. First, the single candidate identification model can be rejected. Recall that in this model we assumed that only the strongest candidate would be identified. In most cases this would be the near candidate. If the near candidate were the distractor, there should have been an effect in the post-target

region as a function of distance. This was not observed. We can also reject a model in which it is assumed that a selection stage occurs during the gaze duration on the post-target region. This would also predict a distractor effect on the post-target region and not the target region.

Two types of models are most consistent with the our results. The first assumes that a compound cue consisting of the adjective-noun is used as a probe to memory. Though the cue has sufficient information to identify a single antecedent, the presence of another exemplar of the noun's category slows the identification process. The use of a unique compound cue eliminates the need for a selection stage. Although this compound cue model cannot be rejected on the basis of our results, the weight of the evidence from previous studies supports a different hypothesis about the activation and further processing of the antecedent. This hypothesis is that there is a selection stage and it occurs during the processing of the anaphoric noun. Let us consider an exhaustive search model of activation and selection that could account for our results.

In the early 80's, Lorch (1982), Ratcliff and McKoon (1981) and Anderson (1983) all independently arrived at similar conclusions involving activation. All three proposed that activation of concepts in memory is a two-stage process, in which several concepts are activated in the first stage and one is selected in the second stage. Lorch concluded that the strength of association between category concepts and their exemplars did not affect the rate of activation but did affect the time required to select among the activated concepts. Anderson cited additional evidence for this position, further arguing that the rate of activation of concepts in the first stage was rapid. Fan effects, processing time increases

due to an increase in the number of items to be processed, were said to be due to a decision process in the second stage. Whereas the Lorch and Anderson conclusions were based on studies of semantic memory, Ratcliff and McKoon (1981) used sentence materials but came to essentially the same conclusion based on the time course of activation spreading between nodes in a paragraph. They also found that the time for activation to reach a node was not a function of distance but occurred very quickly. They concluded that the "amount of activation that arrives at a node falls off as a function of distance, but that the time required for activation to arrive at the node is not a function of distance"(pg.461). Thus, there is sufficient evidence from both semantic (Anderson, 1983; Lorch, 1982) and episodic (Ratcliff and McKoon, 1981) memory studies to believe that the initial identification stage will not be affected by distance, or number of antecedents.

In the second stage, the actual antecedent would need to be selected. If we assume that even when there is no distractor present we need to enter a selection stage, then we can account for a distance effect even in the absence of a distractor. This assumption can be justified because there is always activation of semantic associates even when there is a well learned episodic association (Perlmutter, Harsip, and Myers, 1976). If semantic associates are activated together with the target antecedent, there would need to be selection from among the activated concepts. Assuming that strength decays with time, the relative activation of a near antecedent with respect to its semantic associates would be greater than the relative activation of a far antecedent to its semantic associates. This difference in relative activation would result in an increase in the time to select the antecedent the further back it appears in the text. In addition, if there is a distractor

present, there is an additional competitor in the selection phase. The probability of delay due to an attempt to select a non-antecedent increases when there was a recently processed distractor. This increased probability of a delay that occurs in both the near and far conditions will account for a distractor effect.

Within this exhaustive search model framework we can develop a simple algebraic representation of the time spent in the target region, which we now consider to reflect both identification and selection processes. Given our four conditions and their means in milliseconds per character we have:

	Present	Absent
Near	$N + (1-p)\Delta = 43.59$	$N = 41.81$
Far	$F + p\Delta = 50.42$	$F = 45.35$

where:
F, N are base times/character
$p = P$(near selected first)
$\Delta =$ delay due to selection of wrong antecedent

F and N represent the time of the activation stage plus the time in the selection stage due to distance of the antecedent as explained above; p refers to the probability of being selected first. Δ is the delay associated with selection of the wrong candidate.

At this point, there are several approaches to describing these functions, all of which are consistent with the main effects of distance and distractor on the target region that were observed in Experiment 1. The first is to assume that the observed means are our best estimates of the population means. Solving for Δ we get a value of 6.85 msec/char. Solving for p we get .74 as the probability of the near item selected first. Given

the fact that the near item should have a higher level of activation (Anderson, 1983; Lorch, 1982; Ratcliff and McKoon, 1981) this seems like a reasonable value.

A second approach is to assume two main effects, distance and distractor, and no interaction between the two. This is what the results of the ANOVA indicated. For this to occur, p must equal (1-p) and N<F. The only way for p to equal (1-p) is for p = .5. Given that a more distant item should have a lower lever of activation, this does not seem reasonable.

A third possibility is to look at the simple effects. First, the simple effect of distractor in the near condition (43.59 - 41.81) was not significant. Second, the simple effect of distractor in the far condition (50.42 - 45.35) is significant (F_1=5.349, p=.031), and the simple effect of distance in the distractor present condition (50.42 - 43.59) is significant (F_1=9.772, p=.006). Additionally, the simple effect of distance in the distractor absent condition (45.35 - 41.81) approached significance (p = .11). Duffy and Rayner (1990) did get a significant difference between their near and far positions without a distractor, which leads us to believe that this is a real difference. Setting p = 1.0 will produce this pattern of simple effects. Obviously, p is only approximately 1 or else there would not be any difference between the near-present and near-absent conditions.

From our pattern of simple effects and assuming our observed means are our best estimates of the population means, we can be fairly certain that the value of p is between .74 and 1.0. This result seems reasonable for the proposed exhaustive search model. Furthermore, if we assume that when selection is more difficult, as measured by an increase in time, then there will be a higher probability of additional processing of the

anaphoric region. The fact that the more difficult conditions were the distractor-present/far condition followed by the distractor-present/near conditions nicely accounts for the increase in total time spent in the disambiguating region and probability of regression to the disambiguating region for those two conditions.

One other possible explanation of the results does exist. It may be that the far antecedent is not retrieved at all because its strength is too weak. Suppose that the reader has a time limit for moving the eyes and reads on when that limit is reached. Such a time limit has been suggested by some researchers (Henderson and Ferreira, 1987; although see Kennison and Clifton, 1995). In order to account for the distractor effect, we must also assume that the setting of this criterion is higher when the reader senses that there is activation from more than one episodic source. Note that reading times alone cannot eliminate the possibility that the far antecedent has not actually been retrieved.

In order to assure that the anaphor has been resolved, Experiment 2 was conducted. Experiment 2 used recognition responses to probe words as a measure of what was active in memory after reading the anaphoric reference. This allows us to eliminate the possibility that the far anaphor is not retrieved. Additionally, if there is evidence that the distractor is activated then it would provide support for the conclusion that both distractor and antecedent are identified.

CHAPTER IV

EXPERIMENT 2

A. Introduction

A common technique in the study of anaphor resolution is to use response time to a probe as a measure of activation of antecedents (e.g. Dell, McKoon & Ratcliff, 1983; Gernsbacher, 1989; Green, Gerrig, McKoon & Ratcliff, 1994; Lucas, Tanenhaus, Carlson, 1990). This is usually done by one of three methods, either naming time, lexical decision time or recognition time. While the three measures may reflect slightly different processing, all three appear to be sufficient to test for activation of the probed item in a discourse representation. Experiment 2 tested for activation of the antecedent and distractor by examining recognition response time to the antecedent or distractor after an anaphoric reference in comparison to a baseline recognition time after no anaphoric reference was made. Naming time and lexical decision were not used because of their greater sensitivity to semantic priming from the anaphor to the probe word; this might have masked any difference in activation.

Only the distractor-present passages were used from Experiment 1. In most anaphor resolution experiments, the anaphor has been resolved at least by the end of the sentence if not sooner. Our concern was the activation of candidate antecedents immediately after reading the anaphoric reference when there was more than one candidate present. For this reason, the distractor absent passages were eliminated and probes were presented immediately after reading the anaphoric reference.

33

Gernsbacher's (19893 recognition probe experiments using pronominal anaphors led us to expect that there might be effects of suppression of nonantecedents along with possible facilitation (referred to as enhancement by Gernsbacher) of the antecedent. Suppression refers to a slower response time to a probe word after an anaphoric reference in comparison to a baseline response in the absence of an anaphoric reference. This suppression effect is generally found on the non-antecedent. Facilitation refers to a faster response time to a probe word after an anaphoric reference in comparison to a baseline response in the absence of an anaphoric reference. Though Gernsbacher did not find evidence of facilitation of antecedents after pronominal anaphors, she did find facilitation after repeated name anaphors. McKoon and Ratcliff (1980) also found facilitation in a recognition response task after reading an anaphoric reference. Lucas, Tanenhaus and Carlson (1990) using a lexical decision task found no evidence of facilitation of antecedents but did find suppression of non-antecedents. Although they found facilitation effects for both antecedents and non-antecedents in a naming task, they attributed this facilitation effect to semantic priming.

We have assumed that farther antecedents are more weakly associated with the anaphor and therefore are more slowly selected among the associates activated. This implies that the far antecedent has indeed been identified and that the long gaze duration on the anaphor in Experiment 1 did not just reflect a long search time, with no resolution. Therefore, one goal of Experiment 2 was to test whether there is facilitative priming when the far candidate is the antecedent.

In the model proposed, we have suggested that the selection process has been slowed by the presence of a distractor. If we can show that the distractor has been identified either by a facilitation or a suppression effect on distractor probes, then we could conclude that the selection process has indeed been slowed by the presence of a distractor. The lack of such a result, however, would not be strong enough evidence to conclude that the distractor has not been identified. Our method of having the subject press a button to present the probe word may provide enough time for the antecedent to have been resolved to the point that a short lived facilitation or suppression effect on the irrelevant exemplar may have disappeared (Dell, McKoon & Ratcliff, 1983; Swinney, 1979).

B. Method

1. Participants

Sixty-three University of Massachusetts undergraduates participated in return for course credit. The data from ten participants were dropped because they made more than 33 percent errors in responding to questions presented after the passage. One participant was dropped because of errors on 80 percent of the recognition probe words in the experimental passages.

2. Materials

Participants read the distractor-present passages from Experiment 1. Twelve additional passages using the same format were constructed to increase the number of items in each condition. Both near and far candidates appeared as antecedents and both were probed. Additionally, the line containing the anaphoric reference was altered to a

neutral sentence in each passage to create a no-anaphor reference condition. In the sample passage presented in Table 1, this was done by changing the line "Lucy passed it back through the giant/shady tree to the first boy" to "Lucy made a one-handed catch and threw it back to the first boy." This resulted in 36 passages with six conditions, three types of reference (near, far, or no-anaphor) and two types of probe (near or far). All experimental probe words required 'yes' responses.

The passages were presented one phrase at a time in response to a key press by the participant until the anaphoric reference was read. After it was read, participants were presented with a recognition probe. The recognition probe for the experimental passages was either the antecedent or the distractor from the current passage.

In addition, 40 filler passages were constructed. Three of the filler passages were used as practice before the experiment. The remaining filler passages did not contain anaphoric references and required "no" responses to the probe words. Comprehension questions were presented in a block after every five passages.

3. Design

For each participant the experimental texts were randomly assigned with two constraints: Each participant saw six passages in each condition, and across participants, each passage occurred in each condition an equal number of times. The order of passages remained the same for all participants. Each participant was assigned to one of two material sets. The material sets differed in the order of the two adjective noun candidates in the passages.

4. Procedure

Each participant was run individually in a session that lasted approximately forty-five minutes. All materials were presented on a computer monitor controlled by a microcomputer. Participants were instructed to place their right thumb on an advance key, their right index finger on a "yes" key, and their left index finger on the "no" key. Each passage began with the phrase "press advance key to continue." Participants were instructed to press the advance key when they were ready to begin a passage. With the first press of the advance key the first phrase of the passage was presented at the beginning of a line on the screen. Additional presses presented the next phrase of the passage to the right of the previous phrase until the line was full. The next press of the advance key after the line was full erased the current line and presented the first phrase of the next line where the initial phrase of the previous line had been located. At some point in the passage, after the advance key was pressed instead of the next phrase appearing, the current line was erased and "XXX" appeared two lines above where the current line had been. The "XXX" remained for 500 msec. The "XXX" was immediately erased and replaced with the probe word. Participants were instructed to respond by pressing the "yes" key if the word had appeared in the passage and pressing the "no" key if it did not. After responding to the probe word by pressing either the "yes" key or the "no" key, the last line read returned to the screen. If they probe word was responded to incorrectly, the word "ERROR" appeared on the screen for 1500 msec. before the last line read returned to the screen. If the recognition response was over 1000 msec. then a "too slow" message appeared before the last line read returned. Participants then continued reading to the end

of the passage. Participants were instructed to read at a comfortable pace. Every five passages, the phrase "prepare for test sentences" appeared. This was followed by comprehension questions about the previous set of passages. Participants were instructed to answer the questions by pressing the "yes" or "no" keys. Participants were told that it was important to answer the questions as quickly and accurately as possible. On those trials in which the question was answered incorrectly, the word "ERROR" was presented for 1500 msec. Each session began with three practice passages to make sure that participants understood the procedure.

C. Results

Results are reported for recognition times, reading times and error rates. The means for recognition times are reported in Table 5. Error rates are reported in Table 6. The means for reading times are presented in Table 7. In addition, four passages were deleted from the items analysis of the recognition data; these four items had a missing mean in one of the experimental conditions.

For the recognition data, a 2 (near probe v. far probe) x 3 (near antecedent v. far antecedent v. no antecedent) analysis of variance was performed. For the reading time data, a 2 (near probe v. far probe) x 2 (near antecedent v. far antecedent) analysis of variance was performed. The no antecedent condition was excluded from the reading time analysis because the region was not of the same "determiner-adjective-noun" format as the other two reference types.

1. Recognition Times

Recognition times were faster when a near item was probed than when a far item was probed, $F_1(1,50) = 39.929$, MSE = 10,821; $F_2(1,66) = 13.343$, MSE = 29,089. When the probe was the far antecedent, responses were slowest if the near item was the antecedent; when the probe was near, responses where fastest when the near item was the antecedent. This interaction of probe position and antecedent position was significant, $F_1(2,100) = 4.216$ MSE = 7,387; and marginally so when tested against item variability, $F_2(2,132) = 2.427$, MSE = 18,226, $p < .10$. An important issue is whether the antecedent has been accessed. The results of a 2 x 2 analysis of variance between the means when the antecedent was probed and when there was no anaphoric reference supports the conclusion that the antecedent was accessed. Responses were faster when the antecedent was probed (950+890) than when there was no anaphoric reference (984+916); this facilitation effect was significant, $F_1,(1,50) = 6.639$, MSE = 7014. The interaction between antecedent position and probe position was not significant ($F < 1$). In order to test the hypothesis that the far antecedent is accessed, the means of the far probe-far antecedent condition was compared to the far probe-no antecedent condition (950 v. 984). Response to the far probe when it was the antecedent was found to be faster than responses to the far probe when there was no anaphoric reference. This simple effect contrast was significant, $F_1,(1,50) = 4.544$, MSE = 13,180.

A similar 2 x 2 analysis of variance was performed for conditions in which the distractor was the probed item and the conditions in which there was no anaphoric

reference. The responses to the distractor probes were not significantly different than the response to probes after no anaphoric reference (p > .25).

Table 5

Mean Recognition Times to Probe Words (in msec) as a Function of Antecedent Position and Probe Word Position in Experiment 2.

Probe Position	Antecedent Position					
	Far		Near		No Antecedent	
	Mean	(SD)	Mean	(SD)	Mean	(SD)
Far	950[a]	(196)	1002[b]	(200)	984	(187)
Near	907[b]	(153)	890[a]	(175)	916	(180)

Note. [a]Antecedent is the probe word. [b]Distrcator is the probe word.

2. Reading Times

There were no significant differences in the analysis of the reading time data.

Table 6

Mean Reading Times for Anaphoric Reference Phrases (in msec) as a Function of Antecedent Position and Probe Word Position in Experiment 2.

Probe Position	Antecedent Position			
	Far		Near	
	Mean	(SD)	Mean	(SD)
Far	830[a]	(189)	846[b]	(235)
Near	851[b]	(212)	833[a]	(207)

Note. [a]Antecedent is the probe word. [b]Distrcator is the probe word.

3. Error Rates

Readers were more likely to make incorrect recognition responses when presented with far-probes than when presented with near probes, $F_1(1,51) = 61.674$, MSE = .016; $F_2(1,70) = 23.914$, MSE = .029. In addition, error rates were higher when there was no anaphoric reference than when the antecedent was in either the near or the far position, $F_1(2,102) = 3.128$, MSE = .017; $F_2(1,140) = 2.63$, MSE = .015, p = .076. To further examine this difference, we again contrasted the no anaphoric reference conditions with

41

the conditions in which the antecedent was the probe word. Readers were more likely to respond incorrectly to the probe word when there was no anaphoric reference (21.2+9.0) than when the probed word was the antecedent (14.4+5.4). This difference in error rates was significant, $F1(1,50) = 8.2$, MSE = .017; $F2(1,70) = 5.260$, MSE = .019. Also, as with the recognition times, the comparison was made between the error rate in the far probe-far antecedent condition against the far probe-no antecedent condition. Readers were more likely to respond incorrectly in the no reference condition; $F_1(1,50) = 4.531$, MSE = .052; $F_2(1,70) = 3.567$, MSE = .048, p = .063. The analysis performed on the distractor conditions again did not produce any significant differences (p's > .25).

Table 7

Probability of Incorrect Response to the Probe as a Function of Antecedent Position and Probe Word Position in Experiment 2.

Probe Position	Antecedent Position		
	Far	Near	No Antecedent
	Mean	Mean	Mean
Far	.144[a]	.189[b]	.212
Near	.067[b]	.054[a]	.090

Note. [a]Antecedent is the probe word. [b]Distreator is the probe word.

D. Discussion

The response times from Experiment 2 provided evidence that responses to antecedents are facilitated with respect to a baseline. This result was supported by the error rate data. Errors were less likely when the probe word was the near item. Errors were also less likely to be made when the probe word was the antecedent; this was the same contrast that showed a facilitation effect in the response time data. Responses were also made more quickly and more accurately when the probed item was less distant from the anaphoric reference phrase.

In general, these findings support the conclusion that the antecedent was activated. The facilitation effect and lower error rates for the antecedent in comparison to a baseline would be expected if the antecedent was reactivated. This eliminates the possibility of a hypothesis that the anaphoric reference was only resolved in the near condition and not the far. In addition, the process of activating the antecedent is complete after reading the anaphoric reference. This conclusion is based on the fact that the probes were presented immediately following the anaphoric reference.

If selection had occurred during the processing of the anaphor, as our model predicts, we might have expected some evidence that distractors were activated. Neither facilitation nor suppression effects were found. This contrasts with results of several experimenters (Gernsbacher, 1989; Lucas et al., 1990) who have found that non-antecedents were suppressed. One possible reason that we did not find a significant effect is that categorical noun anaphors are of a less explicit class than repeated name anaphors; if so, there might not have been enough time for suppression effects to occur.

This is consistent with Gernsbacher's (1989) conclusion that pronouns, which are less explicit than repeated name anaphors, triggered suppression more slowly than repeated name anaphors.

Whereas the gaze duration data from Experiment 1 suggest that reading times on the anaphor phrase might be longer when the antecedent was more distant, such effects were not observed. There are several possible reasons for the absence of such effects. First, our gaze duration effect occurred on the target region alone, not on the disambiguating region and the target region together. The combining of these two regions in this experiment may have masked the effect. Second, in Experiment 1 regressions to the disambiguating region were also found to be more likely in the distractor present passages. It is possible that sometimes these regressions occurred after the noun was read but before the phrase advance button was pressed. In such a case, additional variability is added to the reading time measure. Third, it is likely that adding a decision component and a motor response to advance the text will increase the variability in the data. For a discussion of the differences between self-paced reading and eye tracking methods, see Rayner and Pollatsek (1989 p. 184-185).

In summary, the facilitation effect together with the more accurate responses to antecedent probe words support the conclusion that the distance effect obtained in Experiment 1 reflects resolution of the anaphoric antecedent, as opposed to failure to identify the far antecedent. Furthermore, because the probe was placed immediately after the anaphoric phrase, we can conclude that the reactivation of the antecedent occurred during the reading of the anaphoric reference.

CHAPTER V

GENERAL DISCUSSION

In these experiments we investigated the nature of the anaphor resolution process in the presence of a second possible antecedent. In Experiment 1, we found evidence for two conclusions: First, the presence of a distractor slowed the time required to access the antecedent as seen in increased gaze durations on the anaphoric noun. Second, the gaze durations on the target noun reflected both activation and selection processes. Experiment 2 demonstrated readers did indeed access the antecedent; that the increased reading time due to distance and distractor in Experiment 1 did reflect resolution of the antecedent and not simply a long search time.

The results of Experiment 1 are consistent with an exhaustive search model of activation and selection that was based on findings in the memory literature (Anderson, 1983; Lorch, 1982; Ratcliff & McKoon, 1981). This model is an extension of the Duffy and Rayner (1990) two-stage model of anaphor resolution. Within their framework, anaphor resolution consisted of a identification stage and a verification stage. The identification stage occurred during the time the eyes spent on the anaphoric noun. The verification stage was said to occur during the processing of the post-target region. In this extension of the Duffy and Rayner model, the time the eyes gazed at the anaphoric noun reflects time spent in both an identification and selection stage. Based on the Anderson, Lorch, Ratcliff & McKoon conclusions, the identification process is assumed to be rapid and unaffected by number and strength of antecedents in memory. It is the decision

process in the selection stage that is affected by number and strength of associated antecedents.

In order to understand the functioning of the model, let us first consider the case when there is no distractor present. Consistent with the assumption that the identification process is rapid and exhaustive, and the assumption that semantic associates are at least momentarily active (Perlmutter et al., 1976), near and far antecedents are assumed to be quickly activated along with other semantic associates. The far candidate will be weaker because of more time elapsed since it was encountered. It will therefore be less active and have a lower priority in the selection stage. Further, assume that while the far candidate will usually be the most active associate with no distractor present, there will still be some variation as a function of the associative strength of possible competitors, the degree of attention in the initial encoding of the far candidate, etc. When the antecedent is near, these variations will be of less consequence. This accounts for the effect of distance with no distractor present, represented by the N v. F difference in the algebraic model presented in the discussion of Experiment 1. When there is a distractor present, both the distractor and the antecedent (Dell et al., 1983; Kintsch & Mross, 1985; Swinney, 1979) along with semantic associates are briefly activated. This results in a more difficult selection process than when only semantic associates are competing because a recently processed associate is now also competing with the antecedent. Because there is still some variation in typicality even with the high-typical exemplars used, the near candidate will not necessarily always be considered first during the selection phase, but will have a high probability of being selected first. Thus the near candidate has an advantage with or

without a distractor present. This can be seen in the second part of the Near v. Far equation in the distractor present condition as presented in Experiment 1's discussion. Because all antecedents were highly typical exemplars of the anaphor's category, times spent in the verification stage were consistent across conditions in Experiment 1.

Also consistent with the exhaustive search model was the finding that readers noticed the presence of a distractor as indicated by the increase in regressions to, and total time spent in, the disambiguating region in the presence of a distractor. Given that the distractor was also active, regressions to the disambiguating region presumably aided in the selection between the two active concepts.

In Experiment 2, we obtained evidence that the antecedent was active after reading the anaphoric reference. This was seen in a facilitation of recognition responses to the antecedent in comparison to a baseline response to the antecedent in the absence of a anaphoric reference. The resulting lower probability of making an incorrect recognition response when there is an anaphoric reference is also indicative of the antecedent being reactivated. These two results, in combination with the gaze duration data of Experiment 1, support the conclusion that the identification and selection of the antecedent is completed while the eyes remain on the anaphoric noun. Although, there was no indication in Experiment 2 of the distractor being activated, our self-paced procedure may have required a response outside of the critical 250ms. period in which we might have expected a difference (Kintsch & Mross, 1985; Swinney, 1979).

This exhaustive search model of anaphor resolution is in many respects consistent with a general framework of lexical access models found in the lexical ambiguity literature.

The traditional finding is that initially, access of both meanings of a homophone were facilitated over a control word. After approximately 250ms. however, only one of the meanings remained active (Onifer & Swinney, 1981; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979). Thus, it appears that in lexical access a rapid initial "identification" of both possible meanings is completed, and within 250ms. the appropriate meaning is selected.

More recent findings support a reordered-access model (Duffy, Morris & Rayner, 1988), which also assumes that all meanings are accessed, but the relative order of activation is affected by meaning dominance and prior biasing context. Thus, both the exhaustive search model of anaphor resolution's selection mechanism and the reordered-access model function on the basis of relative order of activation. In both models, this relative order is determined by the factors which affect the strength of the item in memory. In the reordered-access model these factors are meaning dominance and prior biasing context (Binder & Morris, 1995; Duffy et al., 1988); in the anaphor resolution model these factors are distance in the text and number of recently encountered potential candidates (distractors). Other textual factors which have been shown to affect strength of an item in memory should also affect relative order of activation, such as the typicality of the exemplar(s) (Corbett, 1984) and elaboration of the candidate antecedents (O'Brien, et al., 1990).

In summary, these experiments demonstrated the importance of determining the affect of multiple antecedents on the anaphor resolution process. By determining that the presence of a distractor had its initial effect while the eyes were on the anaphoric noun, we

were able to further our understanding of the process of anaphor resolution. Specifically, we were able to propose an exhaustive search model that was based on a previously proposed model (Duffy & Rayner, 1990) and with results from the semantic memory literature (Anderson, 1983; Lorch, 1982), from the episodic memory literature (Ratcliff & McKoon, 1981) and from the lexical ambiguity literature (Duffy, Morris & Rayner, 1988; Onifer & Swinney, 1981; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979).

Within the framework of the anaphor resolution process that we have proposed, there are many issues that remain to be explored in future research. Our assumption that the distractor is active in memory before the selection between identified items occurs may be testable. By utilizing an RSVP technique (Dell, McKoon and Ratcliff, 1983; Kintsch & Mross, 1985) it may be possible to probe the readers memory between the identification stage and the selection stage to investigate this assumption. Specifically, an RSVP experiment will enable us to probe memory within the critical period of 250ms. that was found in the lexical ambiguity literature by Swinney (1979) A second issue is whether those factors that have been shown to affect strength of association in memory will affect the relative order of activation in the selection process as we would predict on the basis of our model. Based on Corbett's results (1984) typicality is one such factor. Also, as in the lexical ambiguity literature, prior biasing context, and analogously, elaboration as in the discourse literature should affect the selection stage. Manipulation of these variables should produce patterns of times consistent with the assumptions made in our model. A third issue is whether integrative processes are affected. For example, consider a text in

which a protagonist, Cindy, wants to grab a snack. Cindy has a choice between a ripe apple and a large orange. Cindy then chooses the ripe fruit. If we later refer to the distractor item (the orange), is there an increase in reading time in comparison to a no-reference condition because it is the inappropriate fruit, or is there a decrease in reading time because the distractor had been considered in the identification phase and therefore recently received some activation? In summary, by referring back to either the antecedent or distractor later in the passage, it may be possible to determine if there are long term benefits or costs to considering the distractor but integrating only the antecedent.

Footnotes

[1]First fixation were analyzed for the other two regions but were not reported for two reasons. First, there were no significant effects. Second, first fixations are affected by variables such as landing position in a word which are more likely reflecting ocular motor functioning rather than text integration.

BIBLIOGRAPHY

Anderson, J. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and visual information in reading. Cognitive Psychology, 17, 364-390.

Binder, K.S., & Morris, R.K. (1995). Eye movements and lexical ambiguity resolution: Effects of prior encounter and discourse topic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 1186-1196.

Carroll, P., & Slowiaczek, M.L. (1986). Constraints on semantic priming in reading: A fixation time analysis. Memory & Cognition, 14, 509-522.

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

Clifton, C., & Ferreira, F. (1987). Discourse structure and anaphora: Some experimental results. In M.Colheart (ed.), Attention and Performance 12. London: Erlbaum.

Corbett, A.T. (1984). Pronominal adjectives and the disambiguation of anaphoric nouns. Journal of Verbal Learning and Verbal Behavior, 23, 683-695.

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

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

Duffy, S.A., Morris, R.K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. Journal of Memory and Language, 27, 429-446.

Duffy, S.A., & Rayner, K. (1990). Eye movements and anaphor resolution: Effects of antecedent typicality and distance. Language and Speech, 33(2), 103-119.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. Journal of Verbal Learning and Verbal Behavior, 22, 75-87.

Garrod, S., & Sanford, A.J. (1977). Interpreting anaphoric relations: The integration of semantic information during reading. Journal of Verbal Learning and Verbal Behavior, 16, 77-90.

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

Green, S.B., Gerrig, R.J., McKoon, G., & Ratcliff, R. (1994). Unheralded pronouns and management by common ground. Journal of Memory and Language, 33, 511-526.

Henderson, J.M., & Ferreira, F. (1990). Effects of parafovial processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 417-429.

Kennison, S.M., & Clifton, C., Jr. (1995). Determinants of parafovial preview benefit in high and low working memory capacity readers: Implications for eye movement control. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 68-81.

Kintsch, W. (1988). The role of knowledge in discourse comprehension; A construction-integration model. Psychological Review, 95, 161-182.

Kintsch, W., & Mross, E.F. (1985). Context effects in word recognition. Journal of Memory and Language, 8, 336-344.

Lorch, R.F. (1982). Priming and search processes in semantic memory: A test of three models of spreading activation. Journal of Verbal Learning and Verbal Behavior, 21, 468-492.

Lucas, M.M., Tanenhaus, M.K., & Carlson, G.N. (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. Memory & Cognition, 18, 611-631.

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

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

O'Brien, E.J., Duffy, S.A., & Myers, J.L. (1986). Anaphoric inference during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 346-352.

O'Brien, E.J., & Myers, J.L. (1987). The role of causal connections in the retrieval of text. Memory and Cognition, 15, 419-427.

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

O'Brien, E.J., Shank, D.M., Myers J.L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 410-420.

Onifer, W., & Swinney, D.A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency, meaning and contextual bias. Memory & Cognition, 9, 225-236.

Perlmutter, J., Harsip, J., & Myers, J.L. (1976). The role of semantic knowledge in retrieval from episodic long-term memories: Implications for a model of retrieval. Memory & Cognition, 4, 361-368.

Ratcliff, R. & McKoon, G. (1981). Does activation really spread? Psychological Review, 88, 454-462.

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. Memory & Cognition, 14, 191-201.

Rayner, K., & Pollatsek, A. (1989). The Psychology of Reading. Englewood Cliffs, NJ: Prentice Hall.

Sanford, A.J., Garrod S., Lucas, A., & Henderson, R. (1983). Pronouns without explicit antecedents? Journal of Semantics, 2, 303-318.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledged based processing. Cognitive Psychology, 14, 489-537.

Swinney, D.A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 18, 645-659.

Tananhaus, M.K., Leiman, J.M., & Seidenberg, M.S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. Journal of Verbal Learning and Verbal Behavior, 18, 427-440.

Vonk, W. (1984). Eye movements during comprehension of pronouns. In A.G. Gale and F. Johnson (eds.), <u>Theoretical and applied aspects of eye movement research</u>. Amsterdam: North-Holland.