5-1982

Referential Anaphoric Expressions in Three Tie Locations and their Relationship to Reading Achievement of Eighth Grade Students

Mary Ann Medley
The College at Brockport

Follow this and additional works at: https://digitalcommons.brockport.edu/ehd_theses
Part of the Educational Methods Commons, Language and Literacy Education Commons, and the Secondary Education Commons

To learn more about our programs visit: http://www.brockport.edu/ehd/

Repository Citation
https://digitalcommons.brockport.edu/ehd_theses/1103

This Thesis is brought to you for free and open access by the Education and Human Development at Digital Commons @Brockport. It has been accepted for inclusion in Education and Human Development Master’s Theses by an authorized administrator of Digital Commons @Brockport. For more information, please contact kmyers@brockport.edu.
REFERENTIAL ANAPHORIC EXPRESSIONS IN THREE TIE LOCATIONS
AND THEIR RELATIONSHIP TO READING ACHIEVEMENT
OF EIGHTH GRADE STUDENTS

THESIS

Submitted to the Graduate Committee of the
Department of Curriculum and Instruction
Faculty of Education
State University College at Brockport
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Education

by
Mary Ann Medley
State University College at Brockport
Brockport, New York
May, 1982
SUBMITTED BY:

Mary Ann Medley

Date

APPROVED BY:

Arthur E. Smith
Thesis Advisor

Date

Frances M. Toner
Second Faculty Advisor

Date

Robert B. Gill
Chair, Graduate Policies Committee

Date
REFERENTIAL ANAPHORIC EXPRESSIONS IN THREE TIE LOCATIONS AND THEIR RELATIONSHIP TO READING ACHIEVEMENT OF EIGHTH GRADE SUBJECTS

The purpose of this study was to examine eighth grade subjects' abilities to resolve three types of referential anaphoric expressions across three different tie locations in expository discourse and to determine if the ability to resolve referential anaphoric expressions is significantly related to reading comprehension as measured by either a product-oriented or process-oriented test of reading achievement.

One of the chief means of creating cohesion within a text is through the use of anaphoric expressions which refer a reader back to concepts developed previously. Research at the elementary level has led a number of investigators to conclude that elementary children do not adequately comprehend anaphoric expressions. Other researchers have found that difficulty with pronominal referents extends into the junior-high-school range.

An anaphoric resolution test was developed by the investigator in which each of three types of referential anaphoric expressions were crossed with each of three tie locations, according to classifications established by Halliday and Hasan (1976). A two-way analysis of variance was used to provide insight into the influence that these factors have on the ability of eighth grade subjects to resolve referential anaphora. Scores from this instrument were then correlated with the Gates-MacGinitie Reading Test, Level E, a product-oriented measure of reading achievement, and the New York State Preliminary Competency Test in Reading, a process-
oriented measure of reading achievement.

It was found that the type of referential anaphoric expression did not make a significant difference in their resolution and that there was no significant interaction between the type of referential anaphoric expression and the tie location of the presupposed items. The tie location, however, did make a significant difference in the subjects' abilities to resolve referential anaphora. Further analysis demonstrated that there were significant differences between the immediate and both mediated and remote tie locations, but no significant difference between mediated and remote tie locations.

Correlation coefficients between the anaphoric resolution test and either measure of reading achievement were significant beyond the .01 level of confidence.

It was concluded that with average eighth grade subjects the ability to resolve referential anaphoric expressions in expository discourse is dependent to a significant degree upon the tie locations of the presupposed items when the presuppositions are nouns or noun phrases mentioned explicitly in the discourse. It was significantly easier for subjects to resolve these anaphoric expressions when their presuppositions were in the immediate tie locations than when they were in either the mediated or remote tie locations.

Mean percentages of correct responses were also computed which confirmed results of other studies and led to the conclusion that students in the junior-high-school range also do not adequately comprehend anaphoric expressions.

Further research is needed to: (1) establish the causative
factor in the relationship between anaphoric resolution ability and reading comprehension, (2) determine if instruction in the resolution of anaphoric expressions would increase students' abilities to resolve them, (3) determine if there is a relationship between the ability to resolve anaphoric expressions and measures of intelligence, and (4) determine if other factors involved in anaphoric resolution ability are also significant.

Teachers need to be aware of the difficulty that students have resolving anaphoric expressions and attempt to help them in this area.

Authors and publishers of children's reading material, especially basal reader programs and content area textbooks, should be aware of the fact that popular readability formulas do not take anaphoric resolution ability into account.

Teachers, as well as authors and publishers, should be constantly alert to new research findings on anaphora.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter I</td>
<td></td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Purpose</td>
<td>1</td>
</tr>
<tr>
<td>Questions to be Answered</td>
<td>1</td>
</tr>
<tr>
<td>Need for the Study</td>
<td>2</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>4</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>6</td>
</tr>
<tr>
<td>Summary</td>
<td>7</td>
</tr>
<tr>
<td>Chapter II</td>
<td></td>
</tr>
<tr>
<td>Related Literature</td>
<td>8</td>
</tr>
<tr>
<td>Factors Influencing the Resolution of Anaphora</td>
<td>8</td>
</tr>
<tr>
<td>Memory Search Factors</td>
<td>8</td>
</tr>
<tr>
<td>Factors Involved in Making Inferences</td>
<td>24</td>
</tr>
<tr>
<td>Children's Comprehension of Anaphora</td>
<td>27</td>
</tr>
<tr>
<td>Anaphoric Resolution Ability and Reading Comprehension</td>
<td>50</td>
</tr>
<tr>
<td>Summary</td>
<td>52</td>
</tr>
<tr>
<td>Chapter III</td>
<td></td>
</tr>
<tr>
<td>Design of the Study</td>
<td>56</td>
</tr>
<tr>
<td>Purpose</td>
<td>56</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>56</td>
</tr>
<tr>
<td>Methodology</td>
<td>57</td>
</tr>
<tr>
<td>Subjects</td>
<td>57</td>
</tr>
<tr>
<td>Instruments</td>
<td>58</td>
</tr>
<tr>
<td>Procedures</td>
<td>59</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>61</td>
</tr>
<tr>
<td>Summary</td>
<td>62</td>
</tr>
<tr>
<td>Chapter IV</td>
<td></td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>63</td>
</tr>
<tr>
<td>Purpose</td>
<td>63</td>
</tr>
<tr>
<td>Principal Findings</td>
<td>63</td>
</tr>
<tr>
<td>Additional Findings</td>
<td>66</td>
</tr>
<tr>
<td>Summary</td>
<td>67</td>
</tr>
<tr>
<td>Chapter V</td>
<td></td>
</tr>
<tr>
<td>Conclusions and Implications</td>
<td>69</td>
</tr>
<tr>
<td>Purpose</td>
<td>69</td>
</tr>
<tr>
<td>Conclusions</td>
<td>69</td>
</tr>
<tr>
<td>Implications for Further Research</td>
<td>72</td>
</tr>
<tr>
<td>Implications for Classroom Practice</td>
<td>75</td>
</tr>
<tr>
<td>Implications for Authors and Publishers</td>
<td>76</td>
</tr>
<tr>
<td>Summary</td>
<td>78</td>
</tr>
<tr>
<td>References</td>
<td>82</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

<table>
<thead>
<tr>
<th>Reference Notes</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bibliography</td>
<td>37</td>
</tr>
<tr>
<td>Appendix</td>
<td>89</td>
</tr>
<tr>
<td>A. Anaphoric Resolution Test</td>
<td>92</td>
</tr>
</tbody>
</table>

v
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Analysis of Variance of the Differences in the Resolution of Anaphora Due to the Type of Expression or Tie Location</td>
<td>64</td>
</tr>
<tr>
<td>2. Post Hoc Analysis of Differences Between Means for Tie Locations</td>
<td>65</td>
</tr>
<tr>
<td>3. Correlations Between the Anaphoric Resolution Test and Two Measures of Reading Comprehension</td>
<td>66</td>
</tr>
<tr>
<td>4. Mean Percentages of Correct Responses on the Anaphoric Resolution Test</td>
<td>67</td>
</tr>
</tbody>
</table>
CHAPTER I

Statement of the Problem

Purpose

The purpose of this study was twofold: (a) to examine eighth grade subjects' abilities to resolve three types of referential anaphoric expressions across three different tie locations in expository discourse and (b) to determine if the ability to resolve referential anaphoric expressions is significantly related to reading comprehension as measured by either a product-oriented or process-oriented test of reading achievement.

Questions to be Answered

The study attempted to answer the following questions.

Do the types of referential anaphoric expressions make a significant difference in their resolution in expository discourse?

Do the tie locations of the presupposed items make a significant difference in the resolution of referential anaphoric expressions in expository discourse?

Is there any significant interaction between the types of referential anaphoric expressions and the tie locations of the presupposed items?

Is there a significant correlation between the resolution of referential anaphoric expressions in expository discourse and reading comprehension as measured by a product-oriented measure of reading achievement?

Is there a significant correlation between the resolution of referential anaphoric expressions in expository discourse and reading com-
prehension as measured by a process-oriented measure of reading achievement?

Need for the Study

Psycholinguistic research over the past thirty years has done much to further the understanding of the language process—how language is learned, generated, and comprehended. Much of this attention has been centered on individual sentences.

In the last ten years attention has moved from a focus on how words are strung together to form sentences, to how sentences are strung together to form a text. A text is held together through cohesive ties in meaning. According to Halliday and Hasan (1976), "The concept of cohesion is a semantic one; it refers to relations of meaning that exist within a text, and that define it as a text. Cohesion occurs where the interpretation of some element in the discourse is dependent on that of another" (p. 4).

The most common form in which this cohesion occurs is through anaphoric expressions which refer a reader back to a concept which has gone before. An anaphoric expression in one sentence is dependent for its interpretation on a presupposed item found either directly or indirectly at some earlier location in a text. It may refer the reader back to a concept developed within a single word, clause, or sentence, or even an entire paragraph or chapter of a text.

Beginning in about fourth grade and continuing through high school and beyond, students are expected to gain more and more of their knowledge through the comprehension of the expository discourse found in textbooks, newspapers, periodicals, and reference materials. The amount of referential anaphora found in these materials is great. If readers are to understand
these materials they must be able to resolve the anaphoric expressions that they encounter, for as Nash-Webber (1977) states, "if a reader does not recognize an expression as anaphoric, or if he or she is unable to handle it as the writer intended, then there is no way he or she can build a correct model of the text" (p. 4).

A number of researchers, including Barnitz (1980), Bormuth, Manning, Carr, and Pearson (1970), and Richak (1977) have investigated the ability of elementary school children to resolve anaphoric expressions. They have agreed that children in these grades do not fully comprehend anaphoric expressions.

Other researchers, including Chai (1967) and Loban (1963), have found that difficulties with pronominal referents extend into the junior-high-school range.

Moberly (1979) investigated the tie locations described by Halliday and Hasan (1976) and found that with her fourth and sixth grade subjects the tie location was a significant factor in the students' abilities to resolve anaphoric expressions. Follow-up interviews conducted with her subjects indicated even more than the statistical analysis that the tie location was an important factor.

The taxonomically based tests, which are generally used today to assess reading achievement, typically test the product of reading. Students are asked to read a passage and then answer a series of multiple-choice questions based upon what they have just read.

The tests recently introduced to evaluate reading achievement across New York State have a more process-oriented format. Instead of multiple-choice questions at the end of each passage, these tests contain
blanks where words have been deleted within each passage. The student is asked to select the correct word from five options which are all grammatically correct and semantically plausible in the sentence, if the sentence is considered in isolation. Only one of the options, however, completes a sentence which is semantically plausible in the context of the paragraph in which it is embedded. This format is concerned more directly with the process of reading.

There is some evidence that certain types of anaphora have a significant relationship to general reading comprehension. Results of investigations by Detore (1978) and Dutka (1978) both point to a positive relationship between the anaphora each tested and general reading comprehension. Detore tested third, sixth, ninth, and twelfth grade students using anaphora classified by the types of inference necessary for their resolution. Dutka tested college freshmen using nominal substitution anaphora. Both studies indicated a positive relationship between the anaphora examined and general reading comprehension.

Definition of Terms

The key terms used in this study are defined as follows.

Anaphoric expressions are pronouns and pronoun-like structures which serve to refer a reader back to concepts developed earlier in a text.

Referential anaphora are those anaphoric expressions which, instead of being interpreted semantically in their own right, make reference to something else for their interpretation.

Pronominal anaphora are the personal pronouns when they are used anaphorically.
**Demonstrative anaphora** are the demonstrative adjectives and adverbs and the definite article "the" when they are used anaphorically.

**Comparative anaphora** are the adjectives and adverbs of comparison (e.g., "similar," "equally," "such," and "more") and the comparative forms of ordinary adjectives and adverbs (e.g., "bigger," "fastest," and "more quickly") when they are used anaphorically.

**Presupposed items** are those items to which anaphoric expressions refer. They are often referred to as "antecedents," "referents," or "presuppositions." They are assumed to be in the reader's or listener's mind even though they are sometimes not stated in the context of the discourse.

**Ties** are the semantic relations between anaphoric expressions and their presupposed items.

**Tie locations** are the locations of both the anaphoric expressions and their presupposed items.

**Immediate tie locations** are where the presupposed items are located in the sentences immediately preceding the anaphoric expressions.

**Remote tie locations** are where the presupposed items are located two or more sentences preceding the anaphoric expressions and where the intervening sentence or sentences do not make reference to the presupposed items.

**Mediated tie locations** are where the presupposed items are located two or more sentences preceding the anaphoric expressions and where the intervening sentence or sentences do make reference to the same presupposed items.

**Process-oriented reading tests** are those in which a word is deleted from a sentence within a passage and students must choose from five options
the word which fits the sentence in the context of the surrounding passage.

The Degrees of Reading Power Tests used by New York State are examples of this type of reading achievement test.

Product-oriented reading tests are those in which passages are followed by multiple-choice questions. Students must answer the questions based upon the passage. Most generally used achievement tests, such as the Gates Mac-Ginitie Reading Tests, are of this type.

Limitations of the Study

Only referential anaphoric expressions were examined in this study. Other forms of anaphora, such as substitutions and ellipses, were not considered.

Only those anaphoric expressions were examined whose presupposed items were nouns or noun phrases that were mentioned explicitly in the text. Many anaphoric expressions have as their presuppositions items which can only be inferred from the text. These were not considered. Anaphoric expressions may also have as their presuppositions items which are verbs or verb phrases or concepts developed in entire sentences or longer sections of discourse. These were not examined.

Eighth grade students with a reading achievement level above 8.9 were not included in the study. Students were also not included whose vocabulary achievement scores fell below the 5.0 level. Therefore, any conclusions from this study can not be generalized to include students with very low levels of vocabulary development or students whose comprehension is above the eighth grade level.

Only expository material was considered. Since expository writing is, in a number of respects, different from other forms of written discourse,
results of this study can not be generalized to all forms of discourse. It was only expository writing that was considered.

Summary

The amount of research that has been conducted on anaphora is limited. There is evidence to suggest that students do have difficulty resolving some types of anaphoric expressions. There is also evidence that suggests that junior high school students still have difficulty with pronominal referents. The tie locations have also been shown to affect elementary school children's abilities to resolve some types of anaphoric expressions. Much more research on anaphoric resolution ability is indicated. This study examined: (a) referential anaphoric expressions which had, as their presupposed items, nouns or noun phrases that were mentioned explicitly in expository discourse, (b) the tie locations between the anaphoric expressions and their presupposed items, and (c) the relationship between the resolution of these expressions and reading achievement as measured by two different types of reading tests.
CHAPTER II

Related Literature

The purpose of this study was to examine the abilities of eighth grade subjects to resolve referential anaphoric expressions across three different tie locations. A further purpose of this study was to examine the relationship between the ability to resolve anaphoric expressions and general reading comprehension.

The following review of the literature is concerned with factors influencing the resolution of anaphora, children's comprehension of anaphora, and the relationship between anaphoric resolution ability and reading comprehension.

Factors Influencing the Resolution of Anaphora

Memory Search Factors

Since anaphoric expressions refer a reader back to previous concepts developed within a text, the role that memory plays is important to the resolution of anaphora. In order to understand how anaphor-antecedent relationships are formulated in the reader's mind in the process of reading comprehension, we need to know as much as possible about the factors which influence memory search.

Jarvella (1971), using a listening task and college subjects, demonstrated that accurate recall was best for the last clause of discourse heard and fell off sharply for earlier clauses. He also demonstrated the same effect at sentence boundaries. He concluded that recall is typically quite accurate for only the final sentence of discourse heard and that it is
perfectly verbatim for only the most recent clause. According to Jarvella, these findings suggested that "the processing of discourse is guided by a structural analysis which results in storage of one sentence and one clause at a time" (Jarvella, 1973, p. 428).

Chang (1980) investigated this same clause boundary theory developed and tested auditorily by Jarvella (1971), but he was interested in reading comprehension rather than listening comprehension. He used the 12 sentence pairs developed and tested auditorily by Caplan (1972). The following is an example of such a pair:

Now that artists are working in oil/prints are rare.  
Now that artists are working fewer hours/oil prints are rare.

Each sentence was tested with a probe word and the subject was to indicate whether or not it occurred in the sentence just read. Probe words for all the test sentences were either the last word in the first clause or the first word in the second clause. It was hypothesized that if sentences are processed at clause boundaries it would take less reaction time for subjects to respond to probe words occurring in the final clause than in the initial clause. He also incorporated distractor sentences and natural sentences of different structures that were controlled for the number of words occurring between target (probe) word and the test. Chang's hypothesis was confirmed. The result verified that the most recent clause is in active memory and more readily accessible.

From these investigations it would appear that the surface features of clause and sentence boundaries play a role in memory search. The most recent sentence or the most recent clause of that sentence is the most readily accessible in short-term memory.
Clark and Sengul (1979) looked specifically at anaphoric reference across sentence and clause boundaries. They proposed that there were two possible models of the search process—a "continuity model" and a "discontinuity model."

In the 'continuity model' the entities are laid down like beads on a string. When listeners (readers) try to identify the referent of a noun or pronoun, they search these beads from the final one backward. The 'discontinuity model' is like the 'continuity model' except that it has two strings of beads instead of one. The first, and privileged, string contains the entities mentioned in the current sentence and one sentence back, while the second string contains the entities mentioned two or more sentences back. (Clark and Sengul, 1979, p. 35)

They pointed out that these models should only be construed in a "weak" sense, since other factors also influence the assignment of antecedents for anaphora.

To test these memory search models they developed three-sentence context paragraphs which read like simple descriptions so that there was no intrinsic order to the sentences. This was followed by a test sentence containing either a noun or pronoun referent to an entity developed in one of the three earlier sentences. The three sentences in the context paragraph were counterbalanced so that using the same target sentence, the referent would be in the first sentence in one trial and in the second and third positions in other trials. An example of such a paragraph and its target sentence is as follows:

A broadloom rug in rose and purple colors covered the floor.
Dim light from a small brass lamp cast shadows on the walls.
In one corner of the room was an upholstered chair. (context paragraph)
The chair seemed to be antique. (target sentence)

It was hypothesized that if the "continuity model" is at work then the reaction time for the comprehension of the target sentence "as it fit in with the rest of the story" (p. 35) should be progressively longer the farther
back the antecedent was located. If, however, the "discontinuity model" is in operation then the reaction time for antecedents in the third sentence would be significantly faster than if they were in either sentence one or two.

College undergraduates read the context paragraphs in the upper field of a modified tachistoscope, and then, pressed a button which caused the context paragraph to disappear from the upper field and the target sentence to appear in a lower field. As soon as the subject felt he understood the target sentence he pressed a second button. The reaction time for comprehension of the target sentence could thus be measured in milliseconds. Mean latencies computed from the data clearly demonstrated that the target sentence was comprehended quickly when the referent was mentioned in sentence three of the context paragraph, but equally slowly when it was mentioned in either sentence one or two. Clark and Sengul concluded, then, that the "discontinuity model" is most appropriate as a model of memory search. They also performed further analyses of the latencies, but found that the number of words back was relatively unimportant once the number of sentences back had been accounted for.

Clark and Sengul then performed two further, similar experiments to see whether the discontinuity occurred at sentence boundaries or at clause boundaries. Results of experiment two demonstrated that it was not necessarily the last sentence, but the last major clause that is granted a privileged status in working memory. Experiment three further demonstrated that it was the most recent clause which had the privileged status whether that clause was a major or subordinate clause. After comparing results from all three experiments, they concluded:
These factors—clause boundaries, sentence boundaries, and sheer amount of informational content—may work jointly to compound the effects of clause boundaries on the availability of entities mentioned in discourse. (Clark and Sengul, 1979, p. 40)

Haviland and Clark (1974) proposed a "Given-New Strategy" to explain how information is processed. According to their view, an author syntactically identifies what he considers to be "given" information and similarly identifies what he considers to be "new" information. They then propose that the reader searches memory for a matching antecedent to the 'given' information and on finding it, attaches the 'new' information to the antecedent. If he cannot find a matching antecedent, then he must (a) build some sort of bridging structure, (b) treat all information in the sentence as new and begin construction of a new separate structure, or (c) attempt to recompute what is 'given' and what is 'new' in the sentence. (Haviland and Clark, 1974, p. 518)

To test their theory they hypothesized that when reading sentence-pairs in which the tie between what is syntactically marked as "given" information in the second sentence is directly linked to the first sentence, reaction time for comprehension of the second sentence would be faster than it would be if the tie was either indirect or negative, since it would take time to build the bridging structure (the inference). They constructed sentence pairs such as the following which were labelled "Direct Antecedent Pairs," "Indirect Antecedent Pairs," and "Negative Antecedent Pairs," respectively:

Last Christmas Eugene became absolutely smashed. This Christmas he got very drunk again.

Last Christmas Eugene went to a lot of parties. This Christmas he got very drunk again.

Last Christmas Eugene couldn't stay sober. This Christmas he got very drunk again.
Using college undergraduates as subjects and a modified tachistoscope containing an upper and lower presentation field, they measured in milliseconds the time it took to comprehend the second sentence in each antecedent-anaphor pair and found their hypothesis confirmed. It did, indeed, take longer to comprehend either the "Indirect Antecedent" or "Negative Antecedent" pairs than it did the "Direct Antecedent" pairs. It took time to build the bridging structure, the inference, necessary for comprehension of the "indirect" and "negative" antecedent pairs.

It would appear from Haviland and Clark's (1974) study that there is more to the memory search process than just clause and sentence boundaries. Since the difference between each of the antecedent sentences was more a semantic than syntactic difference, it would appear that semantics is also important in the memory search process. This is not surprising, since many researchers, including Chomsky (1965), Lesgold (1974), McCawley (1976), and Polermo and Molfese (1972), have pointed to the confounding of syntax with semantics.

Chafe (1972) proposed a model of semantic structure which involved a number of semantic constraints, each of which "is essentially a statement that the presence of one item in a semantic structure requires, or entails, the presence of some other item or items" (p. 44). Included in his description were "inherent features" of lexical items which are involved in a person's world knowledge of the lexical items. For example, mention of a bicycle allows the discussion of such things as the frame, the wheels, and the handlebars, since the concept of a bicycle inherently includes these parts. He also pointed to a phenomena he called "foregrounding" by which he referred to "the basic idea that at any one point in a discourse
there are certain concepts which are in the foreground of the minds of the participants in a discourse" (p. 50). According to Chafe, once a particular lexical unit has been mentioned in a text it becomes "foregrounded" and remains for an indeterminant length of time, depending upon such factors as: the number of subsequent sentences in which it is not mentioned, a change of scene, the introduction of new events, and so on, in which the foregrounded items are not involved.

Garrod and Sanford (1977) used a reaction time task to answer the questions of when and how information is integrated into the memory structure from the standpoint of semantics. They based their investigation on the work of Battig and Montague (1969), Rosch (1973), and Wilkins (1971). Battig and Montague developed lists of high-frequency exemplars and low-frequency exemplars of a number of word classes. Rosch and Wilkins demonstrated in separate studies that these high- and low-frequency exemplars differ in the speed with which they are accepted as members of their class. Using sentence pairs, one of which contained a class word and the other which contained either a high- or low-frequency exemplar of that class, Garrod and Sanford demonstrated that the conjoint frequency effect was in operation at the time the second sentence was read. By using further reaction-time situations and other sentences which also included both coreferential and non-coreferential sentence pairs, they concluded that a selective search process is continuously in operation when reading for meaning and that:

if information has just been incorporated into a specific location (in memory) this could be marked as 'open' and subsequently searched for information relating to it. When the topic changes, the location could be closed, and accessed only when the information in a new sentence fails to match that in an open location. (Garrod and Sanford, 1977, p. 89)
They further suggested that "one mechanism whereby information is accepted as relevant is the degree of semantic feature overlap between the instance and the category" (p. 89).

In addition to looking at the clause and sentence boundary factors mentioned earlier, Chang (1980) also studied the role of meaning in active memory. He developed sentences which incorporated compound subjects in the first clause followed by either a repetition of the name or a pronoun (to reinstate the meaning from the first clause) in the second clause. The following four sentences demonstrate the four possibilities for the probe word "John," reinstated name condition, reinstated pronoun condition, control name condition, and control pronoun condition, respectively:

John and Mary went to the grocery store and John bought a quart of milk.

John and Mary went to the grocery store and he bought a quart of milk.

John and Mary went to the grocery store and Mary bought a quart of milk.

John and Mary went to the grocery store and she bought a quart of milk.

It was hypothesized that if the pronoun "he" is as effective at reinstating the meaning of "John" as the word "John," itself, and if subjects use only this meaning information for a response, then responses to the probe word "John" should be just as fast in the reinstated pronoun condition as in the reinstated name condition. If "he" does not reinstate the concept "John" or subjects do not make use of meaning information in active memory, then responses for the reinstated pronoun condition should be no faster than for either of the control conditions. Analysis of variance performed on the results of this experiment demonstrated that there was a
significant main effect for reinstatement. The reinstated name condition was significantly faster than the reinstated pronoun condition, and that the reinstated pronoun condition was significantly faster than the mean of the two control conditions. Chang concluded from these results that both a surface and meaning code are simultaneously at work in short term memory. The faster reaction time for the reinstated pronoun condition over the two control conditions indicated that meaning code is in active memory. Since the reaction time to the reinstated name condition was faster than in the reinstated pronoun condition, the additional surface level information facilitated performance.

Lesgold, Roth, and Curtis (1979) performed a total of five related experiments to test the foregrounded versus backgrounded information suggested by Chafe (1972). They used a tachistoscope and reaction time for the comprehension of a target sentence after college subjects had read context paragraphs to test the following hypotheses:

1. When a sentence refers to earlier material in a discourse, that sentence takes longer to understand if the material to which it refers is backgrounded.
2. Comprehension of a sentence depends not only on its relationship to the immediately preceding sentences, but also involves reinstatement of backgrounded information.
3. Backgrounded information will be reinstated if explicitly mentioned again.
4. Backgrounded information which is not explicitly mentioned again might still be reinstated if concepts are mentioned which were originally processed and linked with that information in the discourse.
5. These reinstatement processes are not limited to comprehension of target sentences, but also occur during passage reading. (Lesgold, et al., 1979, p. 298)

The context paragraphs varied to some extent from one experiment to another depending on the purpose of the individual experiment, but were all similar. The following is an example of the background condition passage
for experiment 2, which was followed by the target sentence "The forest was on fire."

Carol was taking her kids on a weekend camping trip. From the moment she started driving that morning, the kids had been fighting. She liked going camping, but she had to struggle to get there safely. Ahead, she saw the dirt road that led to their campsite. The road was several miles long and went through the densest part of the forest. The forest was so close to the side of the road that there was no place to stop. Carol grew anxious when she noticed that on the other side of the road, a thick cloud of smoke covered the forest. The relative quiet was again broken by her children's piercing squeals at a bee in the back seat. Both of the kids were jumping around, but made no attempt to free the insect. Finally, the bee escaped and they clapped and cheered. Then the children started singing "Old MacDonald" with a bee as one of the animals. (Lesgold, et al., 1979, p. 298)

In experiment 2, in addition to a foreground and background condition, four other conditions were developed by adding a single sentence of varying content to the background condition. The following sentences represent the additional sentences used to develop a complete restatement condition, a subject restatement condition, a context restatement condition, and a non-restatement condition, respectively for the above passage:

Carol continued to be distressed, however, as they drove down the dirt road through the smoked filled section of the forest.

Carol continued to be distressed as she drove through the forest.

Carol continued to be distressed as she drove down the road.

Carol continued to be distressed and lit a cigarette with her lighter. (Lesgold, et al., 1979, p. 298)

The number of subjects varied from 12 to 36 from one experiment to another, but all were University of Pittsburgh undergraduates enrolled in an introductory psychology course. All testing was done individually.

The passage, including any extra sentences relative to different conditions, was presented first on a Datapoint Model 3300 video display terminal connected
to a computer by a serial, 120 character per second line. Subjects pressed a space bar to activate the terminal, read the passage, then pressed the space bar again. This caused the passage to be replaced by the target sentence. Subjects were instructed to press the space bar again when they had read and understood the target. Reading time could thus be measured in milliseconds for both the time it took to read the passage as well as for the time it took to read and comprehend the target.

Results of the five experiments confirmed all of the hypotheses. It was found that foregrounding conditions were faster than backgrounding conditions, that repetition of a target subject after it had been backgrounded resulted in significant facilitation of target comprehension, that mentioning a concept related to the target subject provided almost as much facilitation as repetition of the target subject, itself. Experiment five added indirect foreground and background conditions which were found to be substantially slower than the background condition. Lesgold, Roth, and Curtis concluded that this result verified the findings of Havidland and Clark (1974) and supported Kintsch and Vipond's (1978) suggestion "that inference processes are evoked only after an exhaustive reinstatement search is completed" (p. 302).

In summary Lesgold, et al. stated:

We have shown that there are three ways in which a sentence's meaning can be integrated with that of early sentences of discourse. The antecedent for the given portion of a sentence may be in temporary memory and thus immediately matchable. Or it may be reinstated less quickly when the propositions currently in temporary memory are not directly linked to the antecedent. When these possibilities fail, inferential bridges must be generated by the reader. (Lesgold, et al., 1979, p. 307)

McKoon and Ratcliffe (1980) presented experimental evidence in
support of Chang (1980), Haviland and Clark (1974), and Kintsch and Vipond (1978), which concentrated specifically on inference processing. They proposed a model of simple inference processes which involves three component processes. First, a concept to be inferred has to be accessed in memory. Second, the concept, along with its associated propositions, has to be activated (i.e., brought into working memory). Third, the information responsible for the activation has to be connected to a concept and its associated propositions and then the result stored in long-term memory. (McKoon and Ratcliffe, 1980, p. 680)

To test this model they carried out four experiments. They developed 60 four-sentence paragraphs. In each case the fourth sentence either mentioned a critical word from the first sentence, mentioned a category of the critical word, or mentioned a word unrelated to the critical word. For example, using "burglar" as the critical word, they developed the following three-sentence paragraph with three possible fourth sentences:

1. A burglar surveyed the garage set back from the street.
2. Several milk bottles were piled at the curb.
3. The banker and her husband were on vacation.
4a. The burglar slipped away from the streetlamp.
4b. The criminal slipped away from the streetlamp.
4c. A cat slipped away from the streetlamp. (McKoon and Ratcliffe, 1980, p. 672)

In experiment one, they used 60 similar test paragraphs plus 60 filler paragraphs of various lengths, some of which contained the words tested and some of which did not. A total of 36 Dartmouth undergraduates enrolled in an introductory psychology course were tested individually in a reaction time task. As each sentence of each paragraph appeared on the screen the subject was instructed to read it carefully and then press a space bar for the next sentence. After the bar was pressed following the last sentence the test word appeared on the screen.
Subjects were asked whether or not the test word (the critical word) appeared in the paragraph.

Results demonstrated that presentation of an anaphor served to activate its referent. Response times to the critical word were faster when the final sentence mentioned the category than when it mentioned an unrelated word (with either subjects or materials treated as the random variable). Mean reading times for the final sentence also depended on whether the final sentence mentioned the critical word, the category, or the unrelated word.

An anaphor, according to Kintsch and Vipond (1978), should activate propositions connected with the referent as well as the referent itself. McKoon and Ratcliff (1980) tested this theory in their second experiment by using a replication of experiment one. However, instead of the test word being the referent, the test word was a word connected to the referent in the first sentence (e.g., "garage") that was not repeated in any other sentence of the paragraph. Results of this experiment were exactly as predicted. The mean response times for the critical word and category final sentences were the same and faster than mean response times for the unrelated word final sentences. Thus, it was shown that a concept in the same proposition as the referent is activated by an anaphor.

To test the third part of their theory, that propositions of the anaphor and propositions of the referent are connected together and stored in long-term memory, McKoon and Ratcliff (1980) performed a third experiment using a priming technique. The same 60 experimental paragraphs were used as in the first two experiments. Each trial consisted of two paragraphs which appeared in their entirety on the screen and were each read
and studied for eight seconds before they left the screen. Following the second paragraph, a list of 10 words individually appeared on the screen. Subjects responded to each by pushing "yes" or "no" buttons to indicate whether or not the word had appeared in either of the two paragraphs. Response time for each word was measured.

The list of ten words contained the target word from each paragraph (in the demonstration paragraph, above, the target word was "streetlamp") in a random position in the list. If the word was to be primed by the referent word (e.g., "burglar"), the critical word was placed in the immediately preceding position in the list. If the word was to be unprimed, a word from the other paragraph was placed in the immediately preceding position. A third word (a noun from the first sentence of the paragraph other than the critical word) plus two negative words were also included in the list for each trial of two paragraphs.

It was theorized that the appearance of the critical word immediately before the target should prime the target resulting in a faster reaction time to the target word than in the unprimed condition. Reaction time should be fastest when the target word was primed and the last sentence of the paragraph contained either the critical word or the category of the critical word, since these propositions should be connected in long-term memory.

Results of experiment three demonstrated that reaction time in the primed condition was significantly faster when the final sentence mentioned either the critical word or its category than when it mentioned an unrelated word. Reaction time was slowest when the target was unprimed. Thus, it was demonstrated that propositions from the referent
and the anaphor must be connected in long-term memory.

McKoon and Ratcliffe (1980) undertook still a fourth experiment to rule out a possible alternate interpretation of the results of experiment three. This interpretation would state that the results were due to preexperimental semantic association rather than anaphoric reference. This experiment was run similar to experiment three, but in such a way that the category names (e.g., "criminal") could not be perceived as anaphors. The first sentence of a paragraph was followed by two unrelated sentences from other paragraphs and then the final sentence of the first paragraph that either mentioned the category name or the unrelated word. (The definite article "the" was replaced by the indefinite article "a" so that the final category sentence read as "a criminal..." rather than "The criminal...") The test list used only what would be the primed condition (e.g., "streetlamp" preceded by "burglar"). If "streetlamp" was connected to "burglar" by preexperimental semantic association, then reaction times to the target "streetlamp" should be faster when the final sentence mentioned the category than when it mentioned the unrelated word. Such a result was not found. There was a difference of only five milliseconds between the mean reaction times under the two conditions. It was, therefore, shown that the differences found in experiment three were not the result of preexperimental semantic association.

McKoon and Ratcliffe (1980) concluded that the results of their experiments support their theory of inference processing.

Based upon the above studies, it would appear that many factors are working jointly in the memory search processes which are needed to tie anaphors with their presuppositions in a discourse. Among these are:
(a) the structural distance between the anaphor and its presupposition,
(b) the semantic distance between the anaphor and its presupposition,
(c) whether or not the presupposition is still in short-term memory when
the anaphor is encountered, (d) whether or not the presupposition can be
reinstated into short-term memory when the anaphor is encountered, and
(e) whether or not an inference must be made to connect the anaphor with
its presupposition.

Considering ease of resolution of an anaphor to be indicated by
processing speed, it would appear that resolution of an anaphor would be
easiest when the presupposition is in the immediately preceding clause,
when the semantic distance between the presupposition and anaphor is
short, and when no reinstatement or inference processes are necessary.
Reinstatement of the presupposition into short-term memory involves making
a connection between concepts associated with the anaphor and presupposi-
tion and, therefore, is more time consuming and thus considered more dif-
ficult. The most difficult link to be made in the search process occurs
when the presupposition is not stated explicitly, but must be inferred
and then connected to the anaphor in short-term memory. It would also
appear that in the memory search process a direct connection is attempted
first, a reinstatement process in undertaken next, and an inference pro-
cess is undertaken only when the first two attempts have failed to provide
a presupposition for the anaphor encountered.

Although a number of the above studies tested the same or similar
theories, very little has been done in terms of replications of previous
studies to verify any of the above findings. Much more research is needed
in this area.
Factors Involved in Making Inferences

Nash-Webber (1978) claims that

the use of non-explicit antecedents and referents for anaphoric terms depends upon a contract between speaker and listener. This contract stipulates that if the former uses an anaphoric expression whose antecedent or referent was inferentially derived, the latter both can and will make the same inference. (Nash-Webber, 1978, p. 20)

To explain what guarantees that the same inference will be made by the listener as by the speaker, Nash-Webber points to a discourse model in which she describes discourse entities as

the set of entities 'naturally evoked' by a discourse and linked together by the relations they participate in... The entities 'naturally evoked' by the discourse may have the properties of individuals, sets, stuff, events, activities, etc. (Nash-Webber, 1978, p. 4)

She further demonstrates this by the following context sentence plus alternative subsequent sentences:

Each 3rd-grade girl brought a brick to Wendy's house.
(a) She certainly was surprised.
(b) They knew she would be surprised.
(c) She piled them on the front lawn.
(d) She was surprised that they knew where it was.
(e) Needless to say, it surprised her. (Nash-Webber, 1978, p. 4)

Each of the pronouns in the alternative sentences refers to entities not explicitly stated in the first sentence, but which are called forth in the reader's mind.

Nash-Webber further points to a number of structural rewrite rules which are independent of the content to which they apply and which allow the use of certain inferential anaphora. One such rule is applicable whenever a sentence contains an existentially quantified noun phrase within the scope of a quantified one. A second such rule accounts for conjoined predicates as antecedents where only simple ones have been given
explicitly. A third rule yields class antecedents from the previous mention of a member of that class. An example of each of these is as follows:

Mary gave each girl a T-shirt.
She bought them at Design Research.

I can walk and I can chew gum.
Ford can too but not at the same time.

A German Shepard bit me yesterday.
They are really vicious beasts.

A fourth rule involves what Lakoff and Ross (1972) called "morphologically related" (cited in Nash-Webber, 1977, p. 21) lexical items which yield antecedents for an anaphor such as the following:

John was a guitarist before he lost it on the subway.

Among the many factors which Nash-Webber (1977) cited as relevant to the resolution of anaphora are the following: (a) pronouns are marked for number, gender, and case; (b) general world knowledge, mentioned earlier in relation to Chafe's (1972) work; (c) the Precede-Command Condition described by Wasow (Note 1) as:

A noun phrase A may serve as the antecedent for a pronoun B (which agrees with A in the relevant features, including person, number, and gender) if and only if either
(a) B follows A in the discourse, or
(b) A and B are in the same sentence and B does not command A.
(cited in Nash-Webber, 1977, p. 26)

(d) the notion of "theme" in which the roles of the participants in a discourse are changed as little as possible, (e) the inertial explanation where the syntactic roles of participants are changed no more than is necessary (if a participant is in the subject role in one sentence, the same role will be assigned in subsequent sentences, if this is logical), (f) recency, (g) change of scene, (h) implicit causality which
Garvey, Caramazza, and Yates (1974) believe biases the assignment of an antecedent toward the candidate "primarily responsible for instigating the action or state denoted by the antecedent clause" (cited in Nash-Webber, 1978, p. 26); (i) possible worlds in which an anaphor can refer to a non-existent entity; (j) emphatic stress, which is more pronounced in oral language than written language; and (k) empathy.

A number of researchers have noted that the semantic features of verbs are important in comprehension. Fodor, Garrett, and Bever (1968) purposed that "in general, the greater the variety of deep structure configuration the lexicon associates with the main verb of a sentence the more complicated the sentence should be" (cited in Chomsky, 1969, p. 16). Chomsky found that this appeared to be the case with verbs such as "promise" and "see" when working with young children.

Using ambiguous sentences containing pronouns, Chai (1967) found that the semantic features of transitive verbs influenced antecedent assignments. He found that adult resolution for the pronoun "he" in ambiguous sentences depended, in part, upon whether transitive verbs had an "outward feature" where the action was from the subject to the object (e.g., "strike," "push"), and "inward feature" where the subject received or sensed the action (e.g., "understand," "remember"), or a "reciprocal feature" where the action between the subject and object was reciprocated (e.g., "marry," "meet"). He suggested that "the dictionary for computer processing, using the English language, should contain such a semantic feature for the transitive verbs" (Chai, 1967, p. 127).

Caramazza, Grober, Garvey, and Yates (1977) also pointed to the semantic properties of verbs as important in establishing the referents
for anaphora. They used ambiguous sentences containing subordinate clauses and college undergraduates as subjects and found that "a feature intrinsic to the meaning of many verbs seems to influence the direction of the causal relationship" (Caramazza, et al., 1977, p. 606).

Nash-Webber (1978) points to still another factor involved in the resolution of inferential anaphora. She suggests "that people might vary as to the amount of effort that they will expend inferring an antecedent and how reliable that inference procedure needs to be in order for them to accept its end product as an antecedent" (p. 19).

It would appear, then, that many factors influence the resolution of inferential anaphora. The kind of inference necessary, the semantic features of transitive verbs, as well as the willingness of the reader or listener to find an antecedent and accept it as such. Very little research has been done in the form of experimental evidence in support of these inferential factors.

Much more research is needed in this area by psychologists as well as by linguists, psycholinguists, and educators, since the use of inference involves a great variety of reasoning approaches. Still further research is needed in the area of semantics to understand the role of semantics in anaphoric resolution as well as comprehension in general.

Children's Comprehension of Anaphora

One of the most frequently used forms of anaphora involves the pronoun system. Pronouns not only are used referentially to indicate other participants in a discourse, they also carry with them semantic properties of their own by denoting the number, gender, and case features
of their referents. It is, therefore, fundamental to a study of children's resolution of anaphoric expressions to know at about what age children have acquired knowledge of the pronoun system, including their semantic features.

In 1969 Carol Chomsky reported the results of a study of children's comprehension of a number of linguistic structures. Among other syntactic devices, she tested three structures of pronominalization: (a) the pronoun was in the main clause and preceded a noun phrase (e.g., "He found out that Mickey won the race."), (b) the pronoun was in the subordinate clause and preceded the noun phrase (e.g., "After he got the candy, Mickey left."), and (c) the pronoun was in the subordinate clause and followed the noun phrase (e.g., "Pluto thinks he knows everything."). Eight children each from kindergarten through grade four were tested. Results of the experiment with pronominalization indicated that "the principles of pronominalization appear to be acquired by the majority of children at about the same age," (Chomsky, 1969, p. 109) five or six.

Webster and Ingram (1972, cited in Scholes, 1981) studied the comprehension and production of "he," "she," "him," and "her" in children aged 3.0 to 4.5 years. Sentences involving three levels of complexity were tested: (a) single pronoun in the subject position (e.g., "She is sleeping."), (b) two pronouns in transitive sentences (e.g., "She is pushing him."), and (c) compound sentences (e.g., "The father chases the girl and now he is tickling her."). In a doll-manipulation paradigm, percentages of correct responses from the simplest to the most complex constructions, respectively, were: 88%, 80%, and 79% for children aged 3.0 to 3.5 and 97%, 97%, and 83% for children 4.0 to 4.5 years of age.
Scholes (1981) investigated the ability of children between three and seven years of age to comprehend the semantic components (number, gender, and case) of the personal pronouns. Subjects were presented with arrays of line drawings illustrating the actions described by sentences presented orally to the subjects. Subjects responded by pointing to the appropriate drawing. There were six sentences testing number (e.g., "She is touching him."), six sentences testing case (e.g., "Someone is touching her."), and four sentences testing gender (e.g., "She is touching someone.").

Results showed that comprehension of these forms proceeded in a linear fashion. Performance beyond chance first occurred at five and had reached a 90% level of correct comprehension by age seven. It appeared that the gender constraint was the earliest understood and was followed in a parallel fashion by number and case.

Based upon the above studies it would appear that by the age of seven most children have acquired the basic rules and semantic constraints of the personal pronouns. Studies (to be cited later in this review of the literature) show that in the context of a discourse children have considerably more difficulty with identifying the referents of pronouns.

Another fundamental form of anaphora is the use of the definite article "the." One function of "the" is to signal the fact that a concept about to be mentioned has been mentioned previously and that the reader or listener should have in mind the particular item mentioned earlier. It is used anaphorically.
Warden (1976, 1981) studied the use of the indefinite and definite articles, "a" and "the," by young children. If a concept has been introduced earlier in a discourse, further reference to it should be anaphorically signaled by use of the definite article, "the." The original reference should be signaled by "a" for the purpose of identifying it. Warden tested children from five to eight years of age in both a referent present and referent absent condition and an audience present and audience absent condition. Using pairs of children, one child told another about short films he/she was either watching or had just finished watching (referent present or referent absent condition) while the other child was either in the same room unable to see the film or in a different room connected by telephone (audience present or audience absent condition). Results showed that children of these ages did use "the" correctly as an anaphor, but that they also used it in place of "a" to introduce a referent. They did not use the indefinite article, "a," correctly to a significant degree. Warden concluded that children by these ages do understand the use of the definite article as it is used anaphorically, but that due to either their egocentric failure to see the need to identify their referents for their audience or the complexity of the task demands, the subjects did not consider the rules of article use.

Although more research is needed concerning the use of the definite and indefinite articles, it would appear that children by the age of five do generate structures which use the definite article correctly. (This investigator has found no studies concerned with the comprehension of anaphoric "the" by young children.)
One of the first studies concerned with general anaphoric comprehension abilities of children was conducted by Bormuth, Manning, Carr, and Pearson (1970). They investigated the ability of fourth grade subjects to comprehend three broad classes of literal comprehension skills: sentence comprehension, anaphora comprehension, and inter-sentence comprehension. They tested a total of 25 sentence structures, 14 anaphoric types, and 16 categories of inter-sentence syntactic structures. Each structure was tested by incorporating it in two different sentences or sentence-pairs which were then embedded in a four or five sentence paragraph. All 110 resulting paragraphs were tested by each of four question types. Each of 240 subjects responded to each paragraph with one question type, resulting in about 60 subjects responding to each question type for a given structure. All answers were responded to in short answer format except for the anaphora questions which were in the multiple-choice format.

Results of the analysis of variance performed on the data indicated that for all three classes of comprehension skills the between-structures variance was significant at the .01 level and the interactions was not. The question type was significant at the .01 level only for the sentence structures. It was insignificant for both the anaphoric structures and the inter-sentence structures.

In an effort to formulate a hierarchy of the structures to aid in the teaching of these literal comprehension skills, Bormuth, et al. then ranked each structure from the easiest to the most difficult as determined by the percentage of correct responses. For the anaphoric structures the results indicated that the pro-clause, so (e.g., "Joe may go. If so, we will...") was the easiest to comprehend with 86.8%
correct responses. This was followed by the pro-adverb (e.g., "He works in the cellar. It is cool there.") the relative pronouns (e.g., "The man who lives next door works..."), the pro-verb, so-do (e.g., "Joe likes tennis. So does Bill.") the deleted modifier (e.g., "The small boy came. This boy...") the noun phrase demonstratives (e.g., "The black horse belongs to Joe. That is his...") the numerical pronominal (e.g., "Several men went fishing. Two caught...") the inclusive pronouns (e.g., "Joe, Bill, and Mary went to the show. All enjoyed...") the deleted noun (e.g., "There are red and green apples. The green...are mine.") the pro-verb do-be/have (e.g., "Joe is sick. So is Bill.") the negated pronoun (e.g., "Bill and Joe went shopping. No one bought...") the clause demonstratives (e.g., "Joe is stuck in the mud. This leaves us...") the semantic substitutes (e.g., "Those steel towers are antennas. These objects are...") and the personal pronouns (e.g., "Joe left the room. He had...") which were the most difficult to comprehend with 64.5% correct responses (Bormuth, et al., 1970, p. 354).

Sentence structures and inter-sentence structures were similarly ranked from easiest to most difficult. (For a complete listing of these structures the reader is referred to Bormuth, et al., 1970, pp. 353-354.)

Bormuth et al., were most concerned by the fact that large proportions of the fourth grade subjects tested demonstrated a lack of comprehension "of the most basic syntactic structures by which information is signaled in language" (p. 355). Vocabulary and syntactic complexity was held to a minimum, no time limits were imposed or implied, the questions used were among the most common devices for testing competency, and the questions directly followed the paragraphs, facilitating the
subjects in referring back to the paragraph to answer the questions. As a result, Bormuth, et al. suggested that "these data may actually overestimate the performance of students in actual instructional situations rather than underestimating it" (p. 355).

Lesgold (1974) partially replicated the Bormuth, et al. study of the anaphoric structures in what began as a screening task for an experiment on memory for syntax. Lesgold also tested 14 anaphoric expressions, but only nine were the same structures as those that were used in the previous study. There were a number of other differences between the two studies. This study: (a) used oral, constructed responses, (b) explicitly controlled the number of semantically plausible answers in each passage, (c) counterbalanced the location of the target structures in the passages, and (d) tested 80 third and fourth grade subjects.

Analysis of variance performed on the data from this experiment also demonstrated significant differences among the anaphoric forms tested (p<.0001). However, the ranking of the forms according to percentages of correct responses demonstrated a completely different hierarchy than that demonstrated in the previous study. There was a significant negative correlation between the two sets of mean percentages of correct responses. Lesgold pointed to a number of factors that might be responsible for the differences found: (a) the confounding of syntax with semantics, (b) the extent to which the paragraphs "gave away" the answers, and (c) the amount of processing required to get from a syntactic parsing of a sentence to an underlying cognitive representation (short-term memory factors). He concluded "that the use of difficulty orderings
for syntax without regard to semantics is not likely to lead to improved instruction in comprehension" (p. 338).

Since that time a number of researchers have begun to look at factors which may influence the ease or difficulty children encounter in attempting to resolve anaphoric expressions.

Richek (1976) demonstrated that performance of subjects within a given set of linguistic constraints is variable and dependent on the context of the linguistic structure. She tested third, fourth, and fifth grade subjects by developing twenty sentences which either conformed with or violated the minimal distance principal which states that "the subject of a subordinated complement clause is the noun referent which most nearly precedes that clause" (Rosenbaum, 1967, as cited in Richek, 1976, p. 801). She then constructed a locative statement which in the simple condition followed the test sentence or in the complex condition was embedded in the test sentence as a relative clause.

Simple Condition: Susan convinced Ann to do the homework. Ann was sitting on the chair.

Complex Condition: Susan convinced Ann, who was sitting on the chair, to do the homework. (Richek, 1976, p. 802)

These were then embedded in paragraphs of three additional sentences and followed by a Wh-question asking the subject of the subordinate clause. Results demonstrated that structures violating the MDP (Minimal Distance Principal) have not been fully mastered by children of this age range. She found that insertion of the relative clause was a significant factor as was the violation of the MDP. When these two factors were combined in one sentence, comprehension was seriously impaired. She suggested that these results confirmed that short-term memory plays an important
role in comprehension. She then theorized that the increased semantic processing needed to process the violation of the MDP structures combined with the interruption caused by the insertion of the relative clause, were the main factors causing the experimental results. She concluded that perhaps, as the capacity limit of short-term memory is neared, "processing tends to break down rapidly" (p. 804).

In 1977, Richek continued her investigation of children's comprehension by testing the ability of third grade subjects to process three equivalent anaphoric forms referring back to an explicitly mentioned noun. The three anaphoric forms she tested were: (a) repetition of the noun, (b) use of the pronoun, or (c) the ellipsed form. An example of each is, respectively:

John saw Mary and John said hello to Mary.
John saw Mary and he said hello to her.
John saw Mary and Ø said hello to her. (Richek, 1977, p. 147)

In addition to hypothesizing that there would be a difference in the ability of subjects to process the three forms, Richek also hypothesized that sentence complexity might also be a factor. She therefore, varied the sentence complexity in which all three anaphoric forms were to be tested. The following syntactic factors were tested: (a) either zero or two embedded kernels between the anaphor and its referent (kernel types were either relative clauses, adverbial clauses, complements, or embedded questions), (b) the number of words between the anaphor and its referent (10-12 were considered short and 16-18 were considered long), and (c) parallel or switched sentence construction (Parallel construction occurs when the subject of the first independent clause is also the subject of the second independent clause and
Switched construction occurs when the subject of the first independent clause is the object of the second independent clause. She tested two anaphor-antecedent pairs in conjoined sentences (sentences which contain two independent clauses joined by "and"). The referent was always in the first clause; the anaphor was always in the second clause. Since either anaphor-antecedent pair could be tested in each sentence, the question was also a variable in this study. Because null forms do not occur in switched positions or in the non-subject position in English, the experimental design and resulting data had to be divided into parts—a noun/pronoun portion and a noun/pronoun/null portion.

Results of both sets of data indicated that the paraphrase alternations did significantly affect difficulty. Noun forms were easiest, pronoun forms next easiest, and null forms the most difficult to comprehend. In the analyses of two third grade readers (which was done to form the complexity factors and sentence frames) it was found that children encounter the pronoun and null forms far more frequently than the noun form. In her experiment, Richek found that comprehension of the null form dropped to 60% correct.

Of the four complexity factors—kernels, parallelism, length, and question—only the question factor was significant. This was not as expected, but Richek pointed to a number of reasons why complexity may still be a factor, but one which was masked in this particular design.

First, each sentence frame was used to test each possible alternative anaphoric form so that in the noun/pronoun design there were two sentences constructed from each frame. Sentences from each frame were
controlled for number and type of clauses, but specific words varied. The sentence frame was found to be significant. Richel suggested, therefore, that complexity does make a difference, but this difference was not observable when the overall differences of kernel, length, and parallelism were considered.

Second, there was an enigmatic 3-way interaction among the kernel, length, and parallelism factors in the noun/pronoun analysis. When the number of words between anaphor and antecedent was increased the trend of errors for the kernel and parallelism variables was completely reversed. Richel suggested that in her attempt to locate complexity variables she may have inadvertently confounded variables which affect difficulty.

Third, a post-hoc analysis of the kernel types was investigated. It was found that sentences containing relative clauses correlated positively with errors at the .05 level as did noun complements, while other kernel types were not significant. This would suggest that difficulty might depend on the type of kernel inserted between the anaphor and the antecedent rather than just whether or not a kernel was inserted.

Richel (1977) concluded that anaphoric form plays a significant role in children's comprehension of anaphora and that in further research into the effects of sentence complexity on comprehension, contextual variables should be more finely discriminated.

Barnitz (1980) investigated second, fourth, and sixth grade subjects' comprehension of the pronoun "it" as it occurs anaphorically in a variety of linguistic structures. Three hypotheses were formulated:
Referent type: Test passages where the pronoun refers to a noun or noun phrase will be easier to comprehend than passages in which the pronoun refers to a clause or sentence. Reference order: Passages with forward reference order, where the pronoun follows its referent, will be easier to comprehend than those with backward reference order, where the pronoun precedes its referent. Referent distance: Passages with intra-sentential pronoun-referent structures will be easier to comprehend than passages with inter-sentential pronominal structures. (Barnitz, 1980, p. 271)

Structures, such as the following, were constructed crossing all factors with all other factors:

(1) John and his father wanted to buy a large train set, because it was on sale.
(2) Mary rides her skate board in the busy street, but Marvin does not believe it.
(3) Because it was on sale, John and his father wanted to buy a large train set.
(4) Marvin did not believe it, but Mary rides her skate board in the busy street.
(5) John Boy and Mr. Walton went hunting for the rattlesnake in the woods. Mr. Walton was almost bitten by it.

In contrast to studies previously mentioned (Bormuth, et al., 1970; Lesgold, 1974; Richek, 1976, 1977) subjects in this study were not allowed to look back at a passage to answer the Wh-questions. This was accomplished by having each passage in a test booklet followed by a separate page asking the question for that passage. (The page had to be turned before the question was reached.)

The analysis of variance performed on the data indicated that: grade level was a significant factor (p<.01), referent type was a significant factor (p<.05), and reference order was a significant factor (p<.05). The referent distance variable, however, was not shown to be significant (p>.10).

A second analysis of variance was performed in order to include
Results of this analysis showed that reading ability within each grade level was a significant factor (p<.01). Referent type only approached significance (p<.10) and, as in the previous analysis, referent distance was not significant.

Results of both analyses, therefore, confirmed both the hypothesis concerning referent type and the hypothesis concerning reference order. However, in both analyses the referent distance hypothesis was rejected. Barnitz suggested that this may have been due to the fact that the intra-sentential structures were in many cases inter-clausal for two reasons: (a) backward reference is not possible within the same clause, and (b) distance was sometimes varied arbitrarily to maintain other syntactic elements. He further suggested that the subjects' knowledge of the world may override referent distance (p. 285).

Barnitz also developed a hierarchy of the structures he tested, based upon the mean proportions correct for each passage type. He found that the noun phrase pronominal with forward reference within a sentence, demonstrated by sentence (1) above, was the easiest and was well developed by grade two. The sentential pronouns between sentences were the most difficult, whether the order of the referents was forward or backward, and were problematic for sixth-grade subjects. Only 49% of these structures were correct for the sixth-grade subjects. In general, intra-sentential sentence pronominal structures lagged behind intra-sentential noun phrase structures, but were successfully comprehended by the time children entered sixth grade. Inter-sentential noun phrase pronominal structures were comprehended by sixth grade, but inter-sentential sentence structures were still problematic. Barnitz
suggested that such factors as the subjects' knowledge of the world and peculiarities of passages containing the backward reference structures may have affected the hierarchy. (For a complete listing of the hierarchy by grade level, the reader is referred to Barnitz, 1980, p. 286.)

Barnitz concluded that the study "demonstrated that syntactic aspects of anaphora do contribute to readability...and that syntactic structure plays an important role in children's transition to 'skilled reading'" (p. 237). However, he also cautioned that semantic factors, such as inference and knowledge of the world, also affect the comprehension of syntactic structures, and therefore anaphora.

Dutka (1978) also looked at syntactic factors which might affect the ability to resolve anaphoric expressions. Unlike the other studies, however, she used nominal substitutes rather than pronouns as the anaphoric expressions. Among the syntactic factors tested were: (a) changes in grammatical form of the antecedent, (b) discrepancies between the positions of the anaphor in one layer of a sentence and the position of the antecedent in another layer of the same or a different sentence, (c) the grammatical distance between the anaphor and the antecedent, and (d) the length of the antecedent structure.

Results of her study with college students demonstrated that length was the best single indicator of item difficulty. However, when stepwise multiple regression procedures were used, length and distance combined to form the best indicator of item difficulty. These accounted for about 36% of the total variance.

Dutka concluded from these findings, however, that semantic
factors may be more crucial to anaphora resolution than grammatical relations. She suggested that reference, rather than substitution, might be a factor in comprehension.

As was mentioned in the previous section of this review of the literature, concepts are believed to be organized in a hierarchical network in semantic memory. General concepts are located at the top of this network and more specific concepts are located at successively lower levels. Morles (1978) hypothesized that the farther the pronoun was from the antecedent in the hierarchy the more difficult the anaphor would be to comprehend. She also hypothesized that the higher the anaphor's frequency, the easier it would be to comprehend.

To test these hypotheses, Morles constructed two versions of 30 paragraphs differing in hierarchical distance and anaphor frequency. Sixty-eight fifth-grade subjects responded with short answers to Wh-questions asking for identification of the antecedent of the target anaphora.

Results confirmed both hypotheses—near anaphora were easier to comprehend than far anaphora and high-frequency anaphora were easier to comprehend than low-frequency anaphora. She found that there was no significant interaction between hierarchical distance and pronoun frequency. It would appear from these results that the closer the semantic aspects of the referent to the surface of the hierarchy, the easier it will be to resolve. The farther down the hierarchy the reader must search to find a referent for an anaphor the more difficult that anaphor will be to resolve.

Chai (1967) investigated a number of factors involved in the
resolution of ambiguous English sentences by both children and adults and found that children do not process pronominal referents as adults do. He tested fifth, seventh, and eighth grade students and college sophomores. Using 176 ambiguous sentences each containing the nouns "John" and "Bill" and the pronoun "he," subjects were instructed to mark an "X" on the referent for "he" in each sentence. There were 128 sentences containing real verb-pairs and 48 sentences containing nonsense verb-pairs.

Chai found a number of parameters which appeared to cause a tendency toward random responding. These included: (a) any sentence from the each other type (e.g., "John and Bill Yed each other and/but he Xed him."); (b) syntactically complex sentences involving the passive transformation; (c) all but the identical verb-pairs, which would therefore include the similar verbs (e.g., "hit," "kick"), the logical reversal verbs (e.g., "understand," "answer"), or nonsense verbs; (d) deletion of the word "back" from a sentence; and (e) the younger age of subjects (p. 127). As was mentioned in the previous section of this review of the literature, Chai also found that sentences containing "outward" verb-pairs were easier to resolve than "inward" verb-pairs and that "reciprocal" verb-pairs were the most difficult. Analyses of the fifth grade subjects showed random responses to nearly all sentences. However, by seventh grade analyses showed that the subjects had acquired the semantic constraint imposed by "inward" versus "outward" verbs.

Paris and Lindauer (1976) looked further at inferencing ability of children in recall tasks to see if there was any difference in the amount of inference used by subjects at different grade levels. They
used eight sentences which explicitly or implicitly could state the instrument used to accomplish the action of the verb. For example, "The truckdriver stirred the coffee in his cup (with a spoon)" (p. 219).

In their first experiment, subjects from first, third, and fifth grades were presented four explicit and four implicit sentences which they were asked to repeat and remember. This was followed by a four-minute number-circling task after which subjects were given the instruments as prompts for the recall of each sentence. Results showed that only the fifth-graders could recall the implicitly cued sentences as well as the explicitly cued ones.

The second experiment was a replication of the first, except that subjects were from kindergarten, second and fourth grades and all instruments were implied. Subjects were then tested for recall with either the explicit subject, object, or verb of the sentence or the implied instrument of the action. Results demonstrated that the kindergarten subjects recalled a significantly lower proportion of the implicitly cued sentences than did fourth grade subjects.

Considering the theory proposed by Craik and Lockhart (1972) that there are different levels of sentence processing (cited in Paris and Lindauer, 1976, p. 223) the investigators conducted a third experiment using only first-grade subjects. Each subject was told to remember each of ten sentences, as before, but this time they were told to act out the action since this would help them remember each sentence better. They were given five sentences with explicitly stated instruments and five sentences with implied instruments and then cued by instrument words only. Results demonstrated a dramatic effect, especially compared with
the first graders in experiment one who were tested in the same way with fewer sentences to remember. In experiment one first graders recalled 57.3% of the explicit sentences and 31.3% of the implicit sentences, while in experiment three the first-graders recalled 72% of the explicit sentences and 70% of the implicit sentences. Paris and Lindauer suggested that these results demonstrated that young children can and do infer relationships, but that, at least in this artificial laboratory-type task, they do not automatically process information to a level deep enough to facilitate recall. It was demonstrated that these same children could be "forced" to process information more deeply.

It would appear from this study that children can, as early as the age of five, use inferences in processing language. However, it would appear that they often may not process information to a deep enough level to use inferences or to use inferences as an aid in recall.

Detore (1979) investigated the effect that inference plays on the comprehension of anaphora with third, sixth, ninth, and twelfth grade subjects. She chose ten types of anaphora classified by the type of inference required to resolve them. Passages were selected from science and social studies texts at each grade level which contained examples of each of the anaphora to be tested. Results demonstrated that seven types of inference were significant at the .05 level in one or both contexts at each grade level—subjective, directive, objective, stative, negative, temporal, and instrumental. She suggested that there is a developmental pattern as inferences to resolve anaphora become more complex.

Researchers have also begun to investigate the theory of
constructive memory organization which allows for integration of concepts from more than one sentence. Paris and Carter (1973) tested ten children from second grade and ten children from fifth grade with seven, short, unrelated stories in which inferences could be drawn. The subjects were tested individually. Each subject was told to try to remember each sentence of a three-sentence story because they would be tested on how well they could remember the sentences later. Each sentence of a story was first read to the subject, followed by a five-minute block-sorting task after which the subject was asked if each of four sentences was identical to one heard in the original story. One exact replication sentence and one true inference sentence were included along with two false statements in each test. For example, the following story was used:

The bird is inside the cage.
The cage is under the table.
The bird is yellow. (Paris and Carter, 1973, p. 110)

The sentences below then followed the block-sorting task.

The bird is inside the cage.
The cage is over the table.
The bird is under the table.
The bird is on top of the table. (Paris and Carter, 1973, p. 110)

Results showed there was a strong tendency by subjects of both grade levels to recognize the true inference sentence as being identical to a sentence heard in the original story. Paris and Carter concluded that the results of their experiment indicated that "comprehension is an active, synthetic process primarily dependent upon semantic information" (p. 112).

It would appear from this study that at least by second grade children do use inferences to arrive at integrated representations of
concepts. The question arises, when comparing this study with the Paris and Lindauer (1976) study, whether second grade subjects have automatically learned to process information deeply enough to use inferences in memory organization while first-graders have not yet arrived at this level of automatic processing or whether other factors such as the task demands or the difference in the type of inference tested was responsible for the differing results with the younger subjects.

Lesgold (1972) conducted an experiment concerned with memory integration with children eight to ten years old. He was concerned with whether or not children who had recently learned the semantic constraints of the personal pronouns formed an integrated memory representation of sentences whose propositions were connected by pronouns (e.g., "Randy played drums and he wrote stories."). He conducted one experiment using thirty-one third and fourth grade subjects (pretested for their understanding of pronouns) with eight sentences and a second experiment using 37 different third and fourth grade subjects with 16 sentences.

Subjects were shown each of four sentences individually for ten seconds as the experimenter read each. They were then asked to recall each sentence, or as much of it as they could, when presented with a probe word (either verb, or either object) from a sentence (e.g., "Do you remember anything from the sentence that had...in it?"). Scoring was done by the number of lexical items recalled from the "same" proposition as opposed to the number of lexical items recalled from the "other" proposition. Since there were twice as many items to be recalled from
the "other" proposition as there were from the "same" proposition as the probe word, proportions of "sames" and "others" were computed. When these proportions were analyzed by condition (repeated noun or pronoun conditions) it was found that the condition was not significant, but that the proportion of "sames" was significantly greater \( p < 0.001 \) than the one-third that would be expected if sentences were integrated in memory.

Lesgold concluded that "children about nine years of age do not generally integrate pronoun sentences in memory as adults seem to do" (p. 11) and further that

there is no reason to assume that a special strategy beyond ordinary anaphora comprehension capacity plays a role in determining integration effects in memory. Rather, it appears that anaphora comprehension rules will automatically produce integrated memory representations, as long as \( S \) has learned to represent individual propositions efficiently enough so that they can 'fit' into limited STM simultaneously. (Lesgold, 1972, p. 16)

Halliday and Hasan (1976) developed a theory of textual cohesion in which they stated that "cohesion does not concern what a text means; it concerns how the text is constructed as a semantic edifice" (p. 26).

They then further stated that it

is not a structural relation; hence it is unrestricted by sentence boundaries, and in its most normal form it is simply the presupposition of something that has gone before, whether in the preceding sentence or not. This form of presupposition, pointing back to some previous item, is known as anaphora. (Halliday and Hasan, 1976, p. 14)

They divided cohesive items into categories "based on linguistic form; these are the categories of cohesion that can be recognized in the lexicogrammatical system" (p. 303). The categories are: reference, substitution, ellipses, conjunction, and lexical cohesion. Under each of these categories they listed all the words (or examples of words) that
can be used as cohesive items along with their various functions. (For a complete listing the reader is referred to Halliday and Hasan, 1976, pp. 333-338.)

In addition to classifying cohesive items and their functions, they also considered the placement of these items (usually anaphora) and their presupposed items (antecedents) in a discourse. They referred to the semantic distance between anaphora and their presuppositions as "tie locations" and point to four possibilities: (a) an immediate tie (The referent is located in the immediately preceding sentence.), (b) a remote tie (The referent to an anaphor is located two or more sentences back and the intermediate sentence or sentences do not make reference to the presupposed item.), (c) a mediated tie (The referent is located two or more sentences back and the intermediate sentence or sentences do make reference to the presupposed item.), or (d) a mediated and remote tie (Some intermediate sentences make reference to the presupposed item while others do not.).

It would appear that these "tie locations" would indicate the possibilities of information being "foregrounded" or "backgrounded" which Chafe (1972) referred to in connection with short-term memory. "Immediate ties" would appear to indicate "foregrounded" information. "Remote ties" would appear to indicate that information has become "backgrounded" by the time the anaphor is encountered. "Mediated ties" would appear to indicate that information has been held in the "foregrounded" condition over a section of discourse. The combination of "mediated and remote ties" would appear to indicate that the information may have been "backgrounded" and then possibly reinstated by the time the anaphor is
encountered.

Moberly (1979) tested the abilities of fourth and sixth grade subjects to resolve anaphoric expressions using Halliday and Hasan's (1976) classifications of cohesive items and tie locations. She asked a series of four questions: (a) Is there a difference in fourth and sixth grade subjects' performance on an anaphoric identification task? (b) Is there a pattern of understanding of reference, substitution, ellipses, or lexical anaphora that would suggest a hierarchy of anaphoric form for each grade level? (d) Does the tie location make a difference in the subjects' understanding?

Using a three-way analysis of variance with repeated measures, three significant main effects were found: differences between grade levels was significant at the .05 level, differences among anaphoric types was significant at the .01 level, and differences among tie locations was significant at the .01 level. Further analysis, however, determined that there were significant interactions between anaphoric form and tie locations and between grade, anaphoric form, and tie locations.

Based upon follow-up interviews with subjects, even more than on the statistical analyses, Moberly suggested that the tie location was a significant factor in the children's understanding.

Very little has been done to replicate any of the above studies concerned with children's comprehension of anaphora. The main conclusion to be drawn is that children do not satisfactorily understand anaphoric expressions, even though children as young as seven years of age do appear to understand the semantic constraints of the personal pronoun system.
It appears that the difficulty involved in resolving an anaphoric expression depends to a great extent on the context (both semantic and syntactic) in which the anaphoric expression is found. Among the factors which appear to influence the ability of children to resolve anaphoric expressions are: (a) short-term memory capacity, (b) anaphoric form, (c) tie location, (d) sentence complexity, (e) the type of inference that may be required, (f) the amount of processing necessary to resolve the anaphor, (g) world knowledge, (h) inferencing abilities of the child, (i) the level of processing used by the child, and (j) the child's grade level.

Anaphoric Resolution Ability and Reading Comprehension

To the investigator's knowledge very little research has been done concerning the relationship between the ability to resolve anaphoric expressions and general reading comprehension. A number of studies included in the previous section of this review of the literature (Barnitz, 1980; Chai, 1967; and Moberly, 1979), demonstrated that grade level was a significant factor in the resolution of the anaphoric expressions each of them tested. These results, however, could be due to either a relationship between the age of the subjects and their ability to resolve the anaphora tested, a relationship between the subjects' reading comprehension and the anaphora tested, or a relationship between a combination of both the subjects' ages and reading comprehension and the anaphora tested.

Gottsdaker (1979) used an analysis of covariance to investigate the interference that anaphoric pronouns might have on the reading comprehension of eighth grade subjects. To investigate the interference of
pronouns on comprehension, four cloze passages were constructed in which 0%, 33%, 66%, and 100% of the pronouns were replaced by their referents. Reading comprehension scores were obtained from the Science Research Associates' Achievement Series, Reading Section, Form E, Green Level. Passage form was the independent variable, the raw score from the close tests was the dependent variable, and the SRA reading comprehension score was the covariant. Results demonstrated that there were no significant differences among subjects' cloze scores. However, performance on the 0% cloze passage was lower than on the 100% cloze passage.

It would appear from this study that the use of pronouns does affect the reading comprehension of eighth grade subjects, but that this effect is not significant.

In her study of inferential anaphora with students in grades three, six, nine, and twelve (mentioned in the previous section of this review of the literature), Detore (1978) investigated the relationship between the resolution of inferential anaphora and reading comprehension as measured by the Gates-MacGinitie Reading Tests, Form 2M. She computed a multiple regression with reading comprehension as the dependent variable. The regression demonstrated that objective anaphora was the best predictor of reading achievement and that nine of the ten types of anaphora tested were significant at the .05 level of confidence.

This study demonstrated that inferential anaphora affect reading comprehension, since in the analysis reading comprehension was the dependent variable. However, it is possible that it was strictly the inferential skill that affected comprehension rather than anaphoric resolution ability.
The only study this investigator has found that directly examined the relationship between anaphora and reading comprehension was conducted by Dutka (1978). She investigated the abilities of 172 college students to resolve nominal substitution anaphora and correlated the results of her test of nominal substitution with the Total Reading, Reading Comprehension, and Vocabulary scores obtained from the Diagnostic Reading Tests: Upper Level, Survey Section. Significant positive correlations were found between the substitution scores and all three reading scores. The correlation between total substitution and total reading was .77, between total substitution and reading comprehension was .78, and between total substitution and vocabulary was .68.

Much more research is needed to investigate the relationship between anaphoric resolution ability and general reading comprehension. If such a relationship is found further research needs to be undertaken to determine the causative factor in the relationship. Caution must be taken in any research showing causation because of the many other skills involved in the resolution of specific kinds of anaphora.

Summary

Sentences of a discourse are held together through cohesive ties in meaning. One of the chief means by which cohesion is realized is through the use of anaphoric expressions which refer a reader back to concepts which have gone before.

Because of the very nature of anaphora (that is, backward reference) it is necessary to know as much as possible about how the mind functions to connect an anaphoric expression with its presupposed
item, whether it is explicitly mentioned in the text or implied. We need to know the factors involved in memory search and how the mind searches memory to connect an anaphor with its presupposition.

Research, to date, indicates that when an anaphor is encountered, the mind searches short-term memory for its antecedent. If one is found, the antecedent (and concepts connected with it from the previous discourse) are connected to the anaphor and its concepts and then stored as a unit in long-term memory. Whether or not an antecedent can be directly found depends upon whether the concept has remained "foregrounded," is still in short-term memory, when the anaphor is encountered. This would depend on many factors: including the structural distance between the anaphor and its presupposition and the semantic distance between the anaphor and its presupposition.

Research further indicates that if the presupposition is no longer in short-term memory (has become "backgrounded"), a reinstatement search is undertaken. Propositions which were originally connected to the presupposition and propositions connected with the anaphor may be connected in short-term memory; by an indirect route, the anaphor can be matched with its antecedent.

Only when attempts to find a direct match and a reinstatement match have both met with failure are inferential processes undertaken.

Research indicates that the following are among the many discourse factors which are involved in making inferences: (a) the semantic constraints of the personal pronouns; (b) the semantic features of verbs; (c) the "theme" of the discourse, wherein the roles of the participants are changed as little as possible; (d) the syntactic roles
of the participants of a discourse which are changed as little as possible; 
(e) how recently the participants have been mentioned; and (f) changes of 
scene within a discourse.

It has also been suggested that the reader's willingness to 
search for an antecedent and to accept the end product of the search as 
an antecedent may also be factors involved in inferring an antecedent for 
an anaphor.

By far the most important conclusion to be drawn concerning 
children's comprehension of anaphora is that children do not satisfactor­
ily understand anaphoric expressions. Since anaphora is the primary 
pattern of cohesion to unify a text and to develop semantic relationships 
between concepts developed between sentences, it would appear that the 
inability of students to correctly identify the presuppositions of anaphoric 
items may be a major stumbling block to reading comprehension.

Since research has demonstrated that children by the age of 
seven do understand the semantic constraints of the personal pronoun 
system and the definite article, "the," it would appear that these seman­
tic constraints are not responsible for this inability to resolve anaphoric 
expressions.

Many factors have been shown to influence anaphoric resolution 
ability. Among them are: (a) the grade level of subjects, (b) the 
structural distance between the anaphor and its antecedent, (c) the 
semantic distance between the anaphor and its antecedent, (d) the type of 
anaphoric expression, (e) the amount and type of inference needed to re­ 
solve the anaphor, (f) semantic constraints, (g) syntactic constraints, 
(h) sentence complexity, (i) the reader's knowledge of the world, (j) the
level of sentence processing generated by the reader, and (k) factors involved in the memory search process.

Research has demonstrated that children, as young as six years of age, can and do use inferences to arrive at integrated memory representations of a text. Other research, however, has demonstrated that children may not integrate concepts connected by anaphoric reference as adults seem to do.

A number of attempts to form a hierarchy of anaphoric expressions have been undertaken, but due to the many factors involved in anaphoric resolution, these studies have had conflicting results.

Very little research has been done to investigate the possible relationship between anaphoric resolution ability and general reading comprehension. A number of studies have indicated that there may be such a relationship because of the significance of grade level factors. However, only one study, of nominal substitution anaphora, has investigated a possible relationship with reading comprehension directly. A significant, positive relationship was found in this study.

Much more research is needed in this area: to verify findings of studies already reported, to pinpoint more precisely the factors involved in the resolution of anaphora, and to look more closely at possible relationships between anaphoric resolution ability and reading comprehension.
CHAPTER III

Design of the Study

Purpose

This was a correlational study intended to examine the relationship between referential anaphoric expressions and reading comprehension as measured by two different types of comprehension instruments. It was also intended to examine the resolution of three different types of referential anaphora across three different tie locations in expository discourse when the presupposed items are nouns or noun phrases mentioned explicitly in the text.

Hypotheses

The null hypotheses tested in this study were:

1. The types of referential anaphoric expressions do not make a significant difference in their resolution in expository discourse.

2. The tie locations of the presupposed items do not make a significant difference in the resolution of referential anaphoric expressions in expository discourse.

3. There are no significant interactions between the types of referential anaphoric expressions and the tie locations of the presupposed items.

4. There is no significant correlation between the resolution of referential anaphoric expressions in expository discourse and reading comprehension as measured by a product-oriented measure of reading achievement.

56
5. There is no significant correlation between the resolution of referential anaphoric expressions in expository discourse and reading comprehension as measured by a process-oriented measure of reading achievement.

Methodology

Subjects

This study included thirty-two eighth grade students in a small, semi-rural school district in western New York State. Students who demonstrated a vocabulary achievement level below 5.0 on the vocabulary portion of the Iowa Test of Basic Skills administered in May, 1981, were excluded from the study to minimize the effect that insufficient word recognition might have on the study. Students were also excluded who demonstrated a reading achievement level above 8.9 on the reading portion of the same measure to control for a possible "ceiling effect." The thirty-two students in the study represented the remaining students enrolled in two sections of the eighth grade in the school district. This was a heterogeneous group of average age and intelligence. There were 16 males and 16 females.

Eighth grade students were chosen as subjects for this study for two reasons. First, research has indicated that continued difficulty with pronominal referents extends into the junior-high-school range. Second, the process-oriented measure of reading achievement, the Preliminary Competency Test in Reading, is a state exam which is first administered to students in eighth grade.
Instruments

The instrument used to examine the resolution of referential anaphoric expressions in expository discourse was designed by the investigator as follows.

Two passages were written on topics covered by the eighth grade science curriculum of the participating school, entitled "Violent Storms" and "The Planets." Two passages were also written on topics covered by the eighth grade social studies curriculum. These were entitled "The Civil War" and "Lewis and Clark." A sample passage was also written based on the science curriculum, entitled "Coal." Each passage contained between 300 and 350 words and had a readability level of 5-6 according to the Dale-Chall readability formula.

The referential anaphoric expressions examined in this instrument were pronominals, demonstratives, and comparatives as explained by Halliday and Hasan in *Cohesion in English* (1976). An example of each of these was crossed with the immediate, mediated, and remote tie locations of the presupposed items, also explained by Halliday and Hasan.

Each passage contained one example of each type of referential anaphoric expression crossed with each tie location of the presupposed items. Since there were three anaphoric expressions and three tie locations under study, this yielded a total of nine anaphoric expressions examined in each passage. Only those expressions were examined which had as their presupposed items nouns or noun phrases explicitly mentioned in the passage. Other expressions, although in the passages, were not examined. Sentential presuppositions were not examined in this instrument.
The anaphoric expressions under examination were underlined and numbered by placing an arabic numeral above the underlined expression in the passage. They were numbered chronologically as they occurred in the passages so that the first expression encountered was numbered one (1) and the last expression encountered was numbered nine (9). Directly under each passage were nine Wh-questions asking for the referent of the correspondingly numbered anaphoric expression in the passage. The questions were followed by lined spaces on which subjects recorded their responses.

The passages were arranged in booklet form with the order of passages rotated within test booklets to avoid systematic order affects. In order of occurrence, each test booklet consisted of: a page of directions, a sample passage page, and four pages of the actual test passages. The directions to students, sample passage, and the four test passages are contained in Appendix A.

The New York State Preliminary Competency Test in Reading, Form B, was used as the process-oriented measure of reading achievement. This instrument yields one raw score which was used in the statistical analysis.

The Gates-MacGinitie Reading Test, Level E, Form 3, was used as the product-oriented measure of reading achievement. Only the raw scores from the comprehension section of this instrument were used in the statistical analysis.

Procedures

The test of referential anaphoric expressions was administered by the investigator to small groups of students in a regular classroom
setting during the month of November, 1981. Booklets were distributed and students were instructed to look at the first page, the page of directions. The directions were read orally by the investigator while students read them silently. Students were then instructed to turn to the following page, the sample passage. The investigator read the passage through orally while students read the passage silently. Each question was then read orally and subjects were encouraged to look back to that numbered item in the passage and reread the passage to that point before answering the question. Answers were given orally by individual students for each of the nine questions, in turn, and were recorded by all students in their proper places at the end of the passage. All questions raised by the subjects were answered as they were encountered. When all answers had been recorded and any further questions about procedures had been satisfactorily explained, students were again instructed to read each passage all the way through, to take their time, and to be very specific in their answers. They were also reminded that there were four passages to be completed and that they could have as much time as they needed.

They were then instructed to turn to the first passage and begin.

All answers were scored as either correct or incorrect. There was no partial credit given. All answers which appeared ambiguous in any way were judged by the head of the reading department and the chairman of the secondary English department of the participating school as well as by the investigator. Raw scores used in the analysis of data were the total number of correct answers recorded by each student.

The New York State Preliminary Competency Test in Reading, Form B, was administered to all subjects included in this study plus the
remaining eighth grade students and a few ninth and tenth grade students who needed to take the exam. All procedures were in accordance with those contained in the *Manual for Administrators and Teachers* which accompanies the test. All answer sheets were hand-scored by the examiner. The test was administered during the third week of October, 1981.

The Gates-MacGinitie Reading Test, Level E, Form 3, was administered to small groups of students in a regular classroom setting by the investigator during the last week of September and the first week of October, 1981. Procedures were in accordance with those contained in the *Teacher's Manual* which accompanies the test. Both the vocabulary and comprehension sections of this test were administered. The answer sheets were hand-scored by the investigator, but only the comprehension raw scores were used in the statistical analysis for this study.

**Statistical Analysis**

A two-way analysis of variance was used to determine if the type of anaphoric expression or the tie location of the presupposed items made a significant difference in the subjects' abilities to resolve the expressions and to determine if there was any interaction between these two factors.

A post hoc analysis was then undertaken to determine if there significant differences between the tie locations.

Pearson r correlation coefficients were used to determine any significant relationship between the resolution of referential anaphoric expressions and each of the measures of reading comprehension.

All analyses were based on the raw scores (total number of correct
responses) for each of the measures used in the study.

Summary

An anaphoric resolution test was developed by the investigator in which each of three types of referential anaphoric expressions were crossed with each of three tie locations. A two-way analysis of variance provided insight into the influence that these factors have on the resolution of anaphora. Correlations were also drawn between the anaphoric resolution test and two different measures of general reading comprehension. These provided insight into the relationship between the ability to resolve anaphora and general reading comprehension.
CHAPTER IV

Statistical Analysis

Purpose

The purpose of this study was to examine the ability of eighth grade subjects to resolve referential anaphoric expressions across three different tie locations in expository discourse. A further purpose was to examine the relationship between the ability to resolve anaphoric expressions and general reading comprehension as measured by two different reading comprehension instruments.

Principal Findings

Hypothesis I - The types of referential anaphoric expressions do not make a significant difference in their resolution in expository discourse.

The data failed to reject this hypothesis. As seen in Table 1, the type of referential anaphoric expressions did not make a significant difference in their resolution in expository discourse.
Table 1

Analysis of Variance of the Differences in the Resolution of Anaphora Due to the Type of Expression or Tie Location

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Anaphora</td>
<td>2</td>
<td>62.2</td>
<td>31.1</td>
<td>0.93</td>
</tr>
<tr>
<td>Tie Location</td>
<td>2</td>
<td>351.2</td>
<td>175.6</td>
<td>5.26*</td>
</tr>
<tr>
<td>Tie Location and Anaphora</td>
<td>4</td>
<td>32.7</td>
<td>8.2</td>
<td>0.25</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>902.7</td>
<td>33.4</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>35</td>
<td>1348.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Crit. F = 4.24, *p < 0.05

Hypothesis II - The tie locations of the presupposed items do not make a significant difference in the resolution of referential anaphoric expressions in expository discourse.

The data rejected this hypothesis at the .05 level. The tie location did make a significant difference in the resolution of referential anaphoric expressions. (See Table 1)

A post hoc analysis was then performed to determine which of the three tie locations was significant in the resolution of referential anaphoric discourse (Lindquist, p. 93). As Table 2 indicates, there were significant differences between the immediate tie locations and both the mediated and remote tie locations. There was no significant difference between the mediated and remote tie locations. Anaphoric expressions in the immediate tie locations were significantly easier to resolve than
those in either the mediated or remote tie locations.

Table 2
Post Hoc Analysis of Differences Between Means for Tie Locations

<table>
<thead>
<tr>
<th></th>
<th>Mediated Tie Location</th>
<th>Remote Tie Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate Tie Location</td>
<td>5.34*</td>
<td>7.42*</td>
</tr>
<tr>
<td>Mediated Tie Location</td>
<td></td>
<td>2.08</td>
</tr>
</tbody>
</table>

Crit. d = 4.79, *p<.05

Hypothesis III - There is no significant interaction between the types of referential anaphoric expressions and the tie locations of the presupposed items.

The data failed to reject this hypothesis. There were no significant interactions between the types of referential anaphoric expressions and the tie locations of the presupposed items. (See Table 1)

Hypothesis IV - There is no significant correlation between the resolution of referential anaphoric expressions in expository discourse and reading comprehension as measured by a product-oriented measure of reading achievement.

The data rejected this hypothesis. The Pearson r correlation coefficient between the anaphoric resolution test and the Gates-MacGinitie Reading Test did reach significance. (See Table 3)
Table 3

Correlations Between the Anaphoric Resolution Test and Two Measures of Reading Comprehension

<table>
<thead>
<tr>
<th>Anaphoric Resolution Test</th>
<th>Gates-NagCinitie Reading Test</th>
<th>Preliminary Competency Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.678*</td>
<td>.664*</td>
</tr>
</tbody>
</table>

Crit. $r = .449$. *$p < .01$

Hypothesis V - There is no significant correlation between the resolution of referential anaphoric expressions in expository discourse and reading comprehension as measured by a process-oriented measure of reading achievement.

This hypothesis was also rejected. The Pearson $r$ coefficient between the anaphoric resolution test and the New York State Preliminary Competency Test in Reading did reach a level of significance. (See Table 3)

It may be noted that these correlations were both significant beyond the .01 level ($r = .449$).

Additional Findings

Previous research has led investigators, including Bormuth, Manning, Carr, and Pearson (1970) and Leagold (1974), to conclude that elementary school children do not adequately comprehend anaphoric expressions. Therefore, mean percentages correct were computed from the data to determine if eighth grade subjects also continue to have difficulty in resolving referential anaphoric expressions. Mean percentages of correct responses were found for each cell in the matrix and are listed in Table 4.
Table 4
Mean Percentages of Correct Responses on the Anaphoric Resolution Test

<table>
<thead>
<tr>
<th></th>
<th>Pronominal Anaphora</th>
<th>Demonstrative Anaphora</th>
<th>Comparative Anaphora</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate Tie Location</td>
<td>73</td>
<td>72</td>
<td>63</td>
<td>69</td>
</tr>
<tr>
<td>Mediated Tie Location</td>
<td>59</td>
<td>53</td>
<td>45</td>
<td>53</td>
</tr>
<tr>
<td>Remote Tie Location</td>
<td>52</td>
<td>40</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>55</td>
<td>52</td>
<td>56</td>
</tr>
</tbody>
</table>

Mean percentages of correct responses for each type of anaphoric expression crossed with each tie location ranged from 73% correct for pronominal anaphora in the immediate tie location to 40% correct responses for demonstrative anaphora in the remote tie location. The grand mean percentage of correct responses was found to be 56%. These findings appear to give support to the Bormuth et al. (1970) and Lesgold (1974) studies and to extend these findings to the eighth grade level.

Summary
This study was designed to investigate the ability of eighth grade subjects to resolve three different types of referential anaphoric expressions across three different tie locations. It was further designed to investigate the relationship between the ability to resolve referential anaphoric expressions and general reading comprehension as measured by two different reading comprehension instruments.
It was found that the type of referential anaphoric expression did not make a difference in their resolution and that there was no significant interaction between the type of referential anaphoric expression and the tie location of the presupposed items. The tie location, however, did make a significant difference in the subjects' abilities to resolve referential anaphoric expressions. Further analysis demonstrated that there were significant differences between immediate and both mediated and remote tie locations. Anaphoric expressions in immediate tie locations were significantly easier for the subjects to resolve than those in either mediated or remote tie locations. Immediate and remote tie locations were not significantly different from each other, however.

It was also found that there was a significant correlation between the subjects' abilities to resolve referential anaphoric expressions and general reading comprehension as measured by either a product-oriented or process-oriented measure of reading achievement.

Researchers have concluded from their studies of third and fourth grade subjects that elementary school children do not adequately comprehend anaphoric expressions. Mean percentages of correct responses computed in this study, not only supported their findings, but extended them to the eighth grade level.
CHAPTER V

Conclusions and Implications

Purpose

This study examined the abilities of eighth grade subjects to resolve three types of referential anaphoric expressions across three different tie locations in expository discourse. It further examined the relationship between the ability to resolve these expressions and general reading comprehension. It was hypothesized that an examination of eighth grade subjects' abilities to resolve these expressions would lead to a better understanding of the factors involved in the resolution of referential anaphora and the relationship between this ability and general reading comprehension.

Conclusions

In examining the results of Hypotheses I, II, and III, it was found that there were no significant differences among the types of referential anaphoric expressions and no interactions among the types of expressions and the tie locations of the presupposed items. The tie locations, however, did make a significant difference in the subjects' abilities to resolve the anaphora they encountered. It may be concluded, then, that with average eighth grade subjects the ability to resolve referential anaphoric expressions in expository discourse is dependent upon the tie locations of the presuppositions when the presuppositions are nouns or noun phrases mentioned explicitly in the discourse.

Other researchers have found that the type of anaphoric expression
did make a significant difference in subjects' abilities to resolve them. However, this is not necessarily in conflict with the results found in this study. These investigators were looking at much broader classifications of anaphora. For example, Moberly (1979) examined referential, substitution, ellipses, and lexical anaphora as outlined by Halliday and Hasan (1976). She found a significant difference among these anaphoric forms. The study reported here investigated only the different types of referential anaphora—one category that was examined by Moberly. It may well be that there are differences which are significant among the broader classifications of anaphora, but that any differences which exist among the different types of one category are insignificant.

The fact that the tie location was found to be significant and that there were significant differences between the immediate and either mediated or remote tie locations, but no significant difference between the mediated and remote tie locations appears to be an important finding. Referential anaphora found in an immediate tie location with its presupposed item was significantly easier for eighth grade subjects to resolve than similar anaphora in either mediated or remote tie locations. It would appear that presupposed items which are still "foregrounded," still in short-term memory, are significantly easier to resolve. This finding agrees with results of studies involving both short-term memory research and studies of anaphoric resolution ability.

There was no significant difference between the mediated and remote tie location. There are a number of possible explanations for
this result. Perhaps the repeated reference to a presupposition is not sufficient to hold it in the "foregrounded" condition. Perhaps other factors such as sentence complexity and additional information "overload" the short-term memory capacity of eighth grade subjects and force the referent from short-term memory. Perhaps eighth grade subjects are not aware that the referent of an anaphor needs to be kept in mind to promote understanding. Perhaps other factors, such as the amount of inferencing necessary or general world knowledge caused the results found in this particular study. It might be noted that in looking at the mean percentages of correct responses for tie locations that the anaphora in the mediated tie locations were easier to resolve than those in the remote tie locations. These results were just not statistically significant.

In examining the results of Hypotheses IV and V it was found that the relationship between the ability to resolve referential anaphora and general reading comprehension was significant beyond the .01 level using either a product- or process-oriented measure of reading achievement. These findings would agree with and extend the findings of Dutka (1978) who found a significant relationship between the abilities of college students to resolve nominal substitution anaphora and general reading comprehension. These findings also lend support to the conclusion that the grade level factor found to be significant by Barnitz (1980), Chai (1967), and Moberly (1979) was indeed due to the relationship between anaphoric resolution ability and reading comprehension ability (rather than other factors involved in the grade level factor such as age, general world knowledge, and so on).
It may also be concluded that the relationship which exists between the ability to resolve referential anaphora and general reading comprehension is almost equally measurable by either a process- or product-oriented measure of reading achievement, since the correlation coefficients were so similar (.678 and .664).

In looking at the mean percentages of correct responses, it was found that the range was from 73% for pronominal anaphora in the immediate tie location down to 40% for demonstrative anaphora in the remote tie location with a grand mean of 56%. This supports the findings of other researchers who have concluded that elementary school subjects do not adequately comprehend anaphoric expressions and it extends this conclusion into the junior-high-school range. Since research has indicated that the degree of difficulty subjects encounter with a given anaphoric expression depends to a large extent on the passage in which it is found, these percentages should be looked upon only as a general indication that, at least under some circumstances, eighth grade subjects do have considerable difficulty in resolving anaphoric expressions.

Implications for Further Research

Since the results of this and other studies have pointed to a highly significant relationship between the resolution of anaphora and general reading comprehension, research needs to be conducted to establish the causative factor in this relationship. It would appear that anaphoric resolution ability would be the causative factor, but to this investigator's knowledge, there has been no research conducted to firmly establish causation.
Since repeated studies, including the one reported here, have demonstrated that subjects do not adequately comprehend anaphoric expressions, research needs to be conducted to determine if instruction in the resolution of anaphoric expressions would increase subjects' abilities to resolve them. Would use of cloze techniques increase the abilities of subjects to resolve anaphoric expressions? Would a variety of oral questioning techniques increase the ability of subjects to resolve anaphoric expressions? Would instruction in the use of various anaphoric expressions in written composition classes increase students' abilities to resolve anaphoric expressions while reading? Would instruction in the resolution of anaphoric expressions increase not only anaphoric resolution ability, but also general reading comprehension? Would instruction in the resolution of particular forms of anaphora not only increase students' abilities to resolve them, but also transfer to other forms of anaphora?

Further research also needs to be conducted to determine if there are significant differences among types of anaphoric expressions within other broad classifications of anaphora. Perhaps further research could be conducted using Halliday and Hasan's (1976) classifications to determine if there are significant differences among different types of substitutions or ellipses when the presuppositions are noun or noun phrases explicitly in the discourse. If the results of such studies also demonstrated that differences within each broad category were insignificant, then further research could be conducted to determine if there is a hierarchy of difficulty among the broad classifications. This might
lead to a sequential approach to the teaching of anaphoric comprehension.

Further research also needs to be conducted using other types of presuppositions. Some research has already been conducted showing that sentential presuppositions are more difficult to resolve than noun or noun phrase presuppositions (e.g., Barnitz, 1980), but much more research needs to be conducted in this area.

Further studies also need to be undertaken controlling various syntactic factors. Are anaphora found in simple transitive sentences easier to resolve than those found in passive sentences? Does the addition of relative or subordinate clauses (or other such structures) to sentences affect the difficulty of resolving anaphora in any or all of the different tie locations?

Further research also needs to be done in relation to inferential anaphora. Since ellipses involve at least some degree of inferencing ability, is there a relationship between the ability to resolve this form of anaphora and the comprehension of passive sentences where the instrument (subject) is often omitted and must be inferred? Are there differences among various types of inferential anaphora that affect the ability of students to resolve them?

Further research should also be conducted to see if there is a significant relationship between the ability to resolve anaphora and various intelligence measures. For example, is the ability to resolve anaphora related to the total scores, the verbal or non-verbal portions, or even subscores, of the WISC-R? Perhaps there is a relationship between the digit span subscore of the WISC-R and anaphoric resolution ability
that would point to a relationship between anaphoric resolution ability and short-term memory capacity.

The results of this study supported the research of others demonstrating that elementary students do not adequately comprehend anaphoric expressions and extended this finding to include average students in the junior-high-school range. Would similar results be found with above average eighth grade subjects? Would similar results be found with students in grades nine through twelve? Do students whose reading achievement level is below grade level have more difficulty in resolving anaphora than students of average or above average reading achievement?

This study was based only on the use of expository discourse. Would replications of this study using different passages of expository discourse have similar results? Would replications of this study using other forms of discourse have similar results?

As can be seen much more research is needed on anaphora to understand fully the factors involved in resolving them or to understand how best to develop increased anaphoric understanding in students.

**Implications for Classroom Practice**

Although to this investigator's knowledge no studies have been conducted showing that the ability to resolve anaphora can be enhanced or taught, it would appear that a teacher's awareness of anaphoric relationships and this aspect of the cohesiveness of a text would be an advantage in teaching comprehension. It would appear that helping children to become aware of anaphora and where to look for possible presuppositions would be a worthwhile endeavor.
McCabe (1981) suggests that "Youngsters who are experiencing difficulty in reading often lack what is called 'metatextual' awareness—they are unaware of nuances of language and linguistic conventions" (p. 946). He further recommends that teachers incorporate such questions as "What do you think... means?" or "To what does... refer?" into their teaching of comprehension strategies to develop a sense of "metatextual awareness" in students, to help them become aware of anaphoric relationships.

Perhaps in the not too distant future more specific methods will have been developed and tested through research so that teachers will have a more substantial base on which to teach students about anaphoric relationships and how to resolve them.

Implications for Authors and Publishers

Authors of children's reading material, as well as textbook publishers should be aware of the difficulty that children have in resolving anaphoric expressions. When writing material for children, authors should be aware that the presuppositions for anaphoric expressions may not be available in the minds of their audience. It might be advisable, then, to repeat the referent or even supply a referent (where one is not explicitly mentioned in a text) to promote the reader's understanding.

Many authors and publishers are concerned with the comprehensibility of textbooks and the rewriting of text to reduce the readability level of content material for use by students at lower grade levels or to promote the understanding of students who are underachievers in reading. Durkin (1980) points to the fact that popular readability formulas, such as the
Dale-Chall, the Spache, or the Fry formulas, do not take anaphoric expressions into account. According to the research reported here, the use of additional anaphoric expressions would increase the difficulty of the text for students. Publishers should be aware that replacing difficult words or phrases with anaphoric expressions will often reduce the readability score of a text yet, in reality, will increase the difficulty students will encounter in understanding the passage.

It would further appear from the research reported here that the tie locations between the anaphoric expressions and their presuppositions should be carefully considered. It will be significantly easier for students to resolve an anaphoric expression when the presupposition is in the immediately preceding sentence than when it is located further back in the discourse (whether in the mediated or remote tie location). Careful attention should be made, then, to how far back the presupposition of an anaphor is located. Perhaps it would be better to repeat a referent when it has not been mentioned in the immediately preceding sentence. It would appear that this would be especially true in the expository material used in either science or social studies texts.

Authors and publishers of basal reading programs should also be aware of the difficulty children encounter with anaphoric expressions. Since the use of anaphoric expressions is one of the chief means of promoting cohesion in a text and since the final objective of basal reader programs is to help students become competent, efficient readers, it would appear that the number or even variety of anaphoric expressions should not necessarily be reduced. However, research has shown that
Anaphoric expressions are easiest to resolve when they are in the immediate tie location and when their presuppositions are nouns or noun phrases rather than sentential presuppositions. Further research should lead to a more complete hierarchy of difficulty, but in the mean time, introducing anaphoric expressions with this in mind seems advisable. Noting the introduction of anaphoric expressions in teachers' manuals so that teachers can subsequently discuss their meanings with students also appears advisable.

Authors and publishers, as well as teachers, should be constantly alert to any further research findings on anaphora so that these can be incorporated into basal reader programs, textbook publications, and other children's reading material, as well as incorporated into the teaching of anaphoric understanding in our schools.

Summary

By examining the ability of eighth grade subjects to resolve referential anaphoric expressions across three different tie locations a number of conclusions may be drawn.

The type of referential anaphoric expression did not make a significant difference in the subjects' abilities to resolve them, nor was there a significant interaction between the type of referential anaphoric expression and the tie location of the presupposed items. However, the tie location did make a significant difference. It may be concluded, then, that the ability of eighth grade subjects to resolve referential anaphoric expressions in expository discourse is dependent to a significant degree on the tie locations of the presupposed items.
when these are nouns or noun phrases mentioned explicitly in the discourse. Results further demonstrated that it is significantly easier for eighth grade subjects to resolve referential anaphoric expressions in the immediate tie location than it is those in either the mediated or remote tie locations.

The significant relationship found between the ability to resolve referential anaphoric expressions and two measures of reading achievement supported the results of other studies. It was further found that this relationship was equally measurable by either a product- or process-oriented measure of reading achievement.

Mean percentages of correct responses demonstrated that eighth grade subjects do not adequately understand anaphoric expressions. This supports other studies investigating the ability of elementary school children to resolve various anaphora and extends those findings into the junior-high-school range.

Further research is needed on anaphora to: (1) establish the causative factor in the relationship between anaphoric resolution ability and reading comprehension, (2) determine if instruction in the resolution of anaphoric expressions would lead to an increase in students' abilities to resolve them, (3) determine if there are other significant factors involved in the resolution of various types of anaphora, (4) determine if various types of presuppositions have significant effects on the ability to resolve anaphora, (5) determine if various factors involved in sentence complexity affect the ability to resolve various anaphora, (6) further investigate factors involved in inferential anaphora, and
(7) determine if there is a significant relationship between the ability to resolve anaphoric expressions and various measures of intelligence.

Teachers should be aware of the fact that students do have difficulty resolving anaphoric expressions and should attempt to help them learn to resolve these expressions.

Authors and publishers of children's reading material, especially basal reading programs and content area textbooks, should be aware that popular readability formulas do not take anaphoric resolution ability into account.

Authors and publishers, as well as teachers, should be alert to further research findings concerning anaphoric resolution ability and be prepared to incorporate these findings into their programs.
References


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


Webster, G. O., & Ingram, D. The comprehension and production of the anaphoric pronouns "he, she, him, her" in normal and linguistically deviant children: A preliminary report. *Papers and Reports on Child Language Development* (Stanford University), 4, 55-78.

Reference Notes

BIBLIOGRAPHY
# Bibliography


APPENDIX
Appendix A

Anaphoric Resolution Test

Directions

On the following pages you will find a sample passage and four other passages that might be found in a social studies or science book. You will find that in each passage there are nine underlined words and that each underlined word is numbered. After each passage there is a question asking you what each of the underlined words refers to.

Read each passage all the way through. When you have finished reading the entire passage answer each question about each of the underlined words. To answer each question you will need to look at the passage again and tell exactly what each underlined word refers to. Be very specific in your answers. For example, if the underlined word is "this" and the question is "What is 'this'?" explain exactly what "this" refers to. Do NOT say it refers to "it."

We will do the sample passage together. As we do it you will have an opportunity to ask as many questions as you may have concerning the test. You will then have as much time as you need to do the rest of the passages. Be sure to read each passage all the way through first, take your time, and be very specific in your answers.
All kinds of coal are formed from wood. The chief members of the coal family are peat, lignite, bituminous coal, and anthracite. More plentiful than oil, deposits of the different forms of coal are found the world over.

The process of coal formation begins in swamps. Woody plants slowly decay by a process which leaves most of the carbon in place. After years of such action, peat (a matted mass of twigs, branches, and leaves) is produced. When the water is drained from these places, the peat can be cut into blocks, set out to dry, and then burned as fuel. Drying is important because this material may be 90 percent water. Ireland is known for using this form of fuel.

If peat is allowed to remain where it forms, it slowly changes into a brown coal called lignite. Though more solid than peat, 1/3 it is still soft and crumbly, and is still one-fifth to one-tenth water. It does, however, burn better.

Bituminous coal is by far the most important member of the coal family. It burns more easily than lignite, since it contains much less moisture. It does not crumble as much and, therefore, can be shipped long distances. It also gives off 5/6 more heat. It is formed from lignite by chemical change and pressure in the earth over many thousands or millions of years.

If bituminous coal remains in the earth and is subjected to enough pressure, it slowly changes to anthracite. This is clean to handle and burns with a short, hot, blue flame. Anthracite makes very little smoke, soot, or odor, and burns for the longest time. These are the properties most desired by consumers. However, since it is not as abundant and is harder to mine than bituminous coal, it is more expensive. Coal consumers find that, of all the forms, bituminous coal serves most of their purposes.

1. What places? 
2. What form of fuel? 
3. What is still soft and crumbly? 
4. "Better" than what?
5. "More" heat than what?
6. What is formed from lignite? 
7. What is clean to handle?
8. "Longer" than what?
9. What is not as abundant?
The Planets

There are nine major planets in our solar system. In the order of their distance from the sun, they are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, and Pluto. Mercury is the closest one, while Pluto is the farthest one. All except Neptune and Pluto can be seen from the Earth without a telescope. Neptune is visible with binoculars, but a six-inch telescope is needed to see Pluto. This one is so far away from Earth that it was not discovered until 1930. While, since it is closest to Earth, we know the most about Mars.

Both Mars and Venus were once thought to be inhabited. They are often compared to Earth. Of these, Venus is larger, has a more dense atmosphere and is much hotter. It does not have any moons. Mars has seasons and polar caps like the Earth. It is almost the same size and its day is only about 40 minutes longer. However, its gravity is much weaker and its atmosphere is only about 5 percent as dense. It has two tiny moons, Deimos and Phobos, that are not round like other moons.

In 1964, the Mariner IV satellite took close-up pictures of the surface of Mars showing how very dry it is. It is covered with such things as lava, rocks, and dust. Scientists know that the atmosphere contains no water vapor. However, they believe that the climate of Mars, millions of years ago, must have been much different from what it is today. It must have been warm enough to have running water because there are many channels and other surface features that could only have been made by it. Since the average temperature is -30°F, scientists believe the water has not disappeared, but instead, is frozen in the polar ice caps. (These are larger in summer than during the winter.) There may also be a thick layer of it hidden beneath the surface. Then, Mars must have been a very different planet from the freeze-dried one now observed.

1. "Farthest" from what? ________________________________________

2. What one? __________________________________________________

3. More than what? ____________________________________________

4. Of what? __________________________________________________

5. What does not have moons? __________________________________

6. "As dense" as what? _________________________________________

7. What has two tiny moons? ____________________________________

8. What is hidden? ____________________________________________

9. When? ____________________________________________________
In 1803, with the approval of Congress, Thomas Jefferson bought the Louisiana Territory from France for $15,000,000. He selected Meriwether Lewis and William Clark to explore the new land.

In 1804, Lewis and Clark left St. Louis by boat and started up the Missouri River with 30 men. In late October, they reached the village of the Mandans, who were friendly Indians. Living there with them were a French trapper and his Indian wife, Sacajawea, and their infant son. She belonged to the Shoshone Indians who lived farther west, near the Rocky Mountains. Lewis and Clark hired them as guides. They continued the expedition by boat up the Missouri River until they reached an area in the Rocky Mountains where the river became too narrow for the boats. This land belonged to the Shoshone Indians. One of the Shoshone chiefs was the brother of Sacajawea. With her as interpreter, Lewis and Clark talked with him. He agreed to give them horses and men to help them cross the mountains. Using the horses, they continued across the Rocky Mountains and reached the Columbia River. There, they made canoes and paddled downstream until they reached the Pacific Coast. They arrived there in November, 1805, after travelling through the present states of Idaho, Washington, and Oregon. They built a fort and spent the winter there.

Four months later, Lewis and Clark began their return trip. Lewis went back along almost the same route they had already investigated. Clark took a more southern route so that more territory could be explored. The route followed the Yellowstone River. He explored this river and then met his partner at the head of the Missouri River. They continued on to St. Louis together, arriving there in September, 1806. It had taken them over two years of extremely difficult work to explore the territory.

1. Who were they living with?
2. "Farther" west than what?
3. What land?
4. Who agreed?
5. Where?
6. "Later" than when?
7. "More" southern than what?
8. Who explored this river?
9. What territory?
Violent Storms

There are many kinds of violent storms including tornadoes, cyclones, and hurricanes. Tornadoes are small in size, but very destructive as they sweep down and strike small areas of land in any section of the U.S. Cyclones are huge storms that originate mainly in the temperate climates and sweep across the land. Of all the violent storms, however, hurricanes are probably the most feared. They also sweep over the land, but compared to cyclones, they are much more destructive. They begin in the tropics and contain warmer, more moist air. Their winds are stronger and they deliver more rain. These storms strike the U.S. around the Gulf of Mexico and along the southeast coast. They occur most often in the late summer and early fall. Over the years, they have caused millions of dollars worth of destruction in the form of wind damage and flooding in these areas.

Home and business owners in hurricane regions are more fortunate today than years ago. They can be more prepared. Pilots take weathermen into the centers of hurricanes. They report to forecasters on the ground such things as the wind speeds, the widths, and the directions in which these storms are moving. Now, it is possible to tell where and when one will strike. With this information they can now take precautions to protect their property and their families.

The winds blow in a clockwise circle around a calm center, or an eye, of a hurricane. This is about 5 to 15 miles in diameter. There is little wind here. The winds in an area may be blowing 100 miles an hour as the storm starts through. Then, the eye arrives and the winds become almost calm, blowing no more than 10 to 15 miles an hour. People sometimes believe that the storm is over. After this passes, however, the winds begin to blow with equal force from the opposite direction.

1. "More" rain than what? 

2. What has caused destruction? 

3. What areas? 

4. "More" prepared than what? 

5. Who report to forecasters? 

6. What storms are moving? 

7. Who can take precautions? 

8. Where? 

9. "Equal" to what?
The Civil War

By the start of the Civil War, in 1861, American citizens had already taken part in three wars since the signing of the Declaration of Independence: the Revolutionary War, the War of 1812, and the Mexican War. They had fought the English, the French, and the Mexicans, among others. Their country had grown from the thirteen original colonies to 36 states stretching all the way to the Pacific Ocean. Now, they were preparing for another war, but this war would be different. They had always fought together against someone else, but now they would be fighting against each other. The country was divided—the North against the South. What strengths did each of these have?

There were more people in the North—22,000,000. The South had less than half that number of people, 9,000,000. Of these people, about 3,500,000 were Negro slaves. In general, they continued to work the farms and plantations. The North had more ships and planned to use these to blockade Southern ports. By blockading these ports the Confederates could not receive supplies from foreign countries. It had more factories to manufacture guns, ammunition, and equipment of war than the South. It had more railroads to move soldiers and supplies from place to place. The government of the U.S. was a more experienced government.

It might seem that the North had all the advantages over the South. It had larger armies with superior equipment. But the South did have some advantages, too. There is an advantage in fighting on "home ground." Northern leaders included William T. Sherman and Ulysses S. Grant, but the South, also, had skillful and intelligent leaders. Many Confederate leaders had graduated from West Point. When the war came, they left the United States Army to defend their home states. Among these men were Robert E. Lee and Jefferson Davis.

1. Who were preparing?
2. Which war?
3. "Different" from what?
4. What people?
5. Who continued to work?
6. What had more factories?
7. "More" experienced than what?
8. "Larger" armies than what?
9. What men?