40253

National Library of Canada | Bibliothèque nationale du Canada

CANADIAN THESES ON MICROFICHE | THÈSES CANADIENNES SUR MICROFICHE

NAME OF AUTHOR/*NOM DE L'AUTEUR* Mikell J. Montague

TITLE OF THESIS/*TITRE DE LA THÈSE* The Relationship of Sentence and Contextual Ambiguity to Reading Comprehension

UNIVERSITY/*UNIVERSITÉ* University of Alberta

DEGREE FOR WHICH THESIS WAS PRESENTED/*GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* M Ed.

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE GRADE* 1978

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* W. T. Fagan

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/*DATÉ* 15 Aug./78 SIGNED/*SIGNÉ* Mikell J. Montague

PERMANENT ADDRESS/*RÉSIDENCE FIXE* 9755 - 84 Ave.
Edmonton Alberta T6E 2E8

NL-91 (3-74)

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

THE UNIVERSITY OF ALBERTA

THE RELATIONSHIP OF SENTENCE AND CONTEXTUAL AMBIGUITY TO READING COMPREHENSION

by

© MIKELL J. MONTAGUE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF EDUCATION

DEPARTMENT OF ELEMENTARY EDUCATION

EDMONTON, ALBERTA

FALL, 1978

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled " The Relationship of Sentence and Contextual Ambiguity to Reading Comprehension" submitted by Mikell J. Montague in partial fulfillment of the requirements for the degree of Master of Education.

W. T. Fagan
.....
Supervisor

.....

.....

Date July, 1974

ABSTRACT

Research has identified the ability to cope with ambiguous sentence structures as one measure of linguistic competence. In a 1972 study, Little found that the ability to identify unambiguous and structurally ambiguous sentences through paraphrases of the sentences correlated significantly with reading comprehension ability. He also found significant differences between the students' ability to cope with the three kinds of sentence structures involved, suggesting that different skills, or different levels of the same skill, were involved in dealing with those structures. It was the purpose of this study to determine whether the abilities of identifying structurally ambiguous sentences continued to develop in students from grade five to grade six and from grade six to seven. A second measure of the ability to cope with ambiguous structures was determined to be the skill of disambiguating structurally ambiguous sentences when they were embedded within constraining paragraphs, again through the selection of correct paraphrases. The relationship between these skills and reading comprehension ability was explored, as well as evidence for differences between the abilities to cope with different kinds of structural ambiguities within the tasks of identifying and disambiguating ambiguities.

Identification of ambiguity was measured by the Sentence Interpretation Test, devised by Little (1972); disambiguating ambiguity was measured by the Contextual Ambiguity Test, devised by the investigator; and reading comprehension ability was measured by the Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest. These tests were administered in April, 1973, to thirty students in grade six and thirty students in grade seven who spoke English when they began their schooling, were proficient in reading at least at the grade five level, and were average or above in I.Q.

The data were analyzed through Pearson product-moment correlations and analysis of variance. Results from analysis of the grade seven data were unexpected; and a follow-up interview conducted by the investigator confirmed that the performance of the grade seven group on the two language tests did not reflect actual abilities for the skills measured.

Analysis of the grade six data revealed a significant difference between the tasks of identifying and disambiguating structurally ambiguous sentences, significant differences between the two types of structural ambiguity, surface structure and underlying structure, as students coped with them within the two tasks, and a significant relationship between the abilities measured by the two language tests and

reading comprehension ability. Grade six students also scored [illegible] than grade five students on all aspects of t[illegible]nterpretation Test.

[illegible]uggested that this study be repeated with another grade seven sample and with succeeding grades to determine the relationship between the language skills measured over various grade levels. It was further suggested that research be conducted to determine the effect of direct instruction in the language skills tested on reading comprehension abilities.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

TABLE

LIST OF FIGURES

CHAPTER I

INTRODUCTION

One fact that has been emphasized by research is that reading is both complex and elusive. Among the factors which contribute to the difficulty of defining reading and examining its various processes are the dynamics both of the readers themselves as they develop from initial decoding stages to varying levels of proficiency as mature readers and of the relationship which exists between reading and the written language structures which provide the input. Neither of these factors is readily controlled in research situations. Nonetheless, researchers must continue to pursue them, for they will not have a clear understanding of reading until they understand language itself, especially written language, and understand how children in their various stages of development cope with language structures.

I. THE PROBLEM

While language acquisition begins long before the child reaches school age, the competency with oral language which the beginning student possesses is still far from that of the mature speaker of the language. Furthermore, he must learn to decode and comprehend written language which may or may not contain structures which are identical to the oral forms. Research has begun to focus

on the forms of written language which are found in textbooks in an effort to identify those structures which are potential problem areas for comprehension and which may require specific instructional attention to accelerate reading comprehension abilities before a given level of independence in reading can be expected. Other research has focused on students to gain knowledge of the developmental patterns which they exhibit in the acquisition of various language abilities.

It was the purpose of this study to determine whether one of the previously identified measures of linguistic competence, the ability to identify structural ambiguity in sentences, continues to develop in students from grade five to grade six, and from grade six to grade seven, and to further explore the relationship which appears to exist between this measure of linguistic competency and reading comprehension. Furthermore, an attempt was made to measure the related language skill of successfully disambiguating structurally ambiguous sentences when they were embedded within paragraphs which allowed only one of the possible meanings to hold true. This skill was also examined for signs of developmental acquisition, as well as for its relationship with reading comprehension.

II. DEFINITION OF TERMS

For the purpose of this study, the following terms will be associated with the meaning given in the definitions below.

Written Language is the graphic representation of the English Language as it appears in handwriting or print (Little, 1972).

Ambiguity exists when any stimulus pattern is capable of two or more distinct interpretations. Ambiguities with more than two interpretations will not be used in this study (Little, 1972).

Structural Ambiguity exists when an orthographic form has two distinct phrase markers associated with it in the surface structure (Little, 1972).

Surface Structure Ambiguity is traditionally defined as involving the possibility of two distinct groupings of adjacent words. For example, in the sentence She spoke to the boy with a smile, the prepositional phrase with a smile could function as an adjectival or an adverbial modifier. For the purposes of this study, surface structure ambiguity will be defined as an instance of structural ambiguity in which the ambiguous structure is characteristic of the ambiguous structures found in sentences traditionally classified as containing surface structure ambiguity and as identified by MacKay and Bever (1967) and

Jurgens (1971). These structures are described in Appendix B (Little, 1972).

Underlying Structure Ambiguity is traditionally defined as involving a change in the logical relations between words rather than a change in the apparent grouping of words. For example, in the sentence The choice of the students was announced, the noun phrase the choice of the students is seen as originating in either of two underlying structures: either the students chose someone (or something) or someone chose the students. In either case the grouping of the words in the surface structure (without labelled bracketing) remains invariant. For the purposes of this study, underlying structure ambiguity will be defined as an instance of structural ambiguity in which the ambiguous structure is characteristic of the ambiguous structures found in sentences traditionally classified as containing underlying structure ambiguity and as identified by McKay and Bever (1967) and Jurgens (1971). These structures are described in Appendix B (Little, 1972).

Identification of Ambiguity is the ability to determine whether a sentence is in fact structurally ambiguous by correctly classifying paraphrased meanings of a structurally ambiguous or unambiguous sentence as

giving a meaning of that sentence (as in the Sentence Interpretation Test) (Little, 1972).

Recognition of Ambiguity is the ability to orally describe the two meanings of a structurally ambiguous sentence or to orally affirm that two meanings are possible when appropriate constraining contexts are provided for the two meanings.

Contextual Paragraph will refer to the paragraph into which an ambiguous sentence was embedded which constrained the meaning of that ambiguous sentence so that only one of the two possible meanings remained true.

Disambiguate is the ability to select the one paraphrase of of an ambiguous sentence which is suitable for the contextual paragraph in which it has been embedded (as in the Contextual Ambiguity Test).

Linguistic Competence refers to one's knowledge of grammar (implicit or explicit) as reflected in his performance in a language situation. In this study, it refers to the ability to utilize specific language structures as conveyors of information.

Reading Comprehension Ability will refer to the child's score on the Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest. This ability can be divided into literal comprehension ability and inferential comprehension ability on the basis of this test score (Little, 1972).

Literal Reading Comprehension Ability is the ability to understand information that is contained explicitly in the material read and in this study will refer to the score on those items in the Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest which the authors of this test claim to measure literal comprehension ability (Little, 1972).

Inferential Reading Comprehension Ability is the ability to understand information that is contained implicitly in the material read and in this study will refer to the score on those items in the Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest which the authors of this test claim to measure inferential comprehension ability (Little, 1972).

Intelligence Quotient will refer to the student's score on on the Canadian Multi-Level Edition of the Lorge-Thorndike Intelligence Test, Levels D or E, for grades six and seven respectively.

III. HYPOTHESES

From the findings of research studies and in view of what the investigator proposes to do in this study, the following research and null hypotheses have been formulated:

Research Hypothesis I

Grade seven students will achieve higher scores than grade six students on

a) Sentence Interpretation Test

b) Contextual Ambiguity Test

Null Hypothesis I

There is no significant difference between the scores of grade six students and grade seven students on

a) Sentence Interpretation Test

b) Contextual Ambiguity Test

Research Hypothesis II

Students in both grades six and seven will score higher on the Contextual Ambiguity Test than on the ambiguous portions of the Sentence Interpretation Test.

Null Hypothesis II

There is no significant difference between the scores of students in both grade six and grade seven on the Contextual Ambiguity Test and the ambiguous portion of the Sentence Interpretation Test.

Research Hypothesis III

Students who score higher on the structural ambiguity tests will also score higher on the reading comprehension test in both grade six and grade seven.

Null Hypothesis III

There is no significant relationship between scores on a test of reading comprehension (Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest) and

scores on a test of identification of ambiguity (Sentence Interpretation Test) or scores on a test on the ability to disambiguate structurally ambiguous sentences in context (Contextual Ambiguity Test) for students in grade six and for students in grade seven.

Research Hypothesis IV

High readers in both grade six and grade seven will be better able than low readers to

a) identify ambiguity

b) disambiguate ambiguous sentences in context

Null Hypothesis IV

There is no significant difference between high and low readers in grade six and grade seven on their ability to

a) identify ambiguity

b) disambiguate structurally ambiguous sentences in context.

Research Hypothesis V

The I.Q. and chronological age of a student in either grade six or grade seven will relate to his ability to identify ambiguity.

Null Hypothesis V

There is no significant relationship between the scores for students in grade six or students in grade seven on the Sentence Interpretation Test and

a) I.Q.

b) chronological age

Research Hypothesis VI

The I.Q. and the chronological age of a student in either grade six or grade seven will relate to his ability to disambiguate ambiguous sentences in context.

Null Hypothesis VI

There is no significant relationship between the scores for students in grade six or students in grade seven on the Contextual Ambiguity Test and

a) I.Q.

b) chronological age

Research Hypothesis VII

Girls in both grade six and grade seven will be better able than boys to

a) identify ambiguity

b) disambiguate ambiguous sentences in context

Null Hypothesis VII

There is no significant difference between boys and girls in grade six and grade seven in their ability to

a) identify ambiguity

b) disambiguate ambiguous sentences in context

Research Hypothesis VIII

Students in grade six and grade seven will be better able to identify unambiguous sentences than structurally ambiguous sentences, and they will be better able to identify sentences with surface structure ambiguity than sentences with underlying structure ambiguity.

Null Hypothesis VIII

There is no significant difference for students in grade six or grade seven among scores on unambiguous sentences, sentences with surface structure ambiguity, and sentences with underlying structure ambiguity on the Sentence Interpretation Test.

Research Hypothesis IX

Students in grade six and grade seven will be better able to disambiguate sentences with surface structure ambiguity than sentences with underlying structure ambiguity.

Null Hypothesis IX

There is no significant difference for students in grade six and students in grade seven between scores on passages containing surface structure ambiguity and passages containing underlying structure ambiguity on the Contextual Ambiguity Test.

The level of significance set for the rejection of the null hypothesis is .05.

IV. ASSUMPTIONS

It is assumed that there are two types of ambiguity. The descriptions of surface structure ambiguity and underlying structure ambiguity which are utilized in this study grew out of the Chomskyan tradition of transformational grammar which includes the concepts of deep structure and surface structure as part of the syntactic component of a

generative grammar. Other linguists, however, have questioned the deep structure concept (Rommetveit, 1968; Prideaux, 1972; Harris, 1970). The interpretation of the findings of this study do not depend on adherence to the deep structure concept as formulated by Chomsky (1965). The differences in difficulty of the two types of ambiguity have been shown to exist by Jurgens (1971) and Little (1972), which appears to indicate they are discreet linguistic structures in terms of how native speakers of English operate on them. For the purposes of this study, the terms were retained in order to be consistent with previous studies.

V. LIMITATIONS

The generalizability of the findings of this study are limited in accordance with the following considerations:

1. The population from which the sample was chosen was limited to grade six and grade seven students in one school of the Edmonton Catholic School System No. 7.
2. Only those students whose scores on the school-administered Gates-MacGinitie Reading Test indicated that they were reading at the grade five level or above were selected for study.
3. Only students who were of average or above intelligence were selected for study.

VI. SIGNIFICANCE OF THE STUDY

Explorations of the relationships between various language factors and reading comprehension have suggested that specific teaching of the ways in which some language structures signal meaning could improve the reading comprehension of school age children. The present study will provide additional information on the relationship between the linguistic competency which is represented by the ability to identify ambiguity in sentences and the ability to disambiguate ambiguous sentences which are embedded in paragraphs and measure reading comprehension achievement. This information will contribute to the knowledge of the researcher, the reading clinician, and the classroom teacher, who are concerned with enhancing the reading comprehension process.

The findings of the study will suggest an expected level of attainment for the measured linguistic skill through grade seven. This information can be used diagnostically by the teacher and the reading clinician by providing an area of focus for raising linguistic competency and subsequent reading comprehension levels or vice versa. The language tests which are used in the study may be used as, or may lead to the development of, useful diagnostic tools for determining linguistic competency.

VII. OVERVIEW OF THE STUDY

The remaining chapters will be divided between the various aspects of the study according to the following scheme. Chapter II will consist of the theoretical framework under which this study was conducted and will review the empirical and theoretical research pertinent to the problem. Chapter III will contain the research design of the study with descriptions of the sample, the experimental and standardized tests used, the pilot study, and the collection and methods of analyzing the data. The results of the test data for the grade six students will be analyzed and explained in Chapter IV. The results of the test data for the grade seven students will be analyzed and explained in Chapter V, as will the findings of a follow-up interview conducted with some of the subjects. Chapter VI will contain the summary, conclusions, and implications.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter consists of an overview of the literature which points toward a need for further exploration of the effect of ambiguous written structures on reading comprehension. Under the heading of "Reading Comprehension and Language Factors", the research reviews have been grouped as to those primarily concerned with readability, those exploring language factors which could affect readability, and those which explore language factors which could affect reading comprehension. Following that, those studies are reviewed which focus specifically on the concept of ambiguity as defined by the transformational-generative grammarians.

Reading Comprehension and Language Factors

Although reading researchers have been concerned for more than fifty years with the problems of how children understand what they read, the reality of classroom practices reflects a lack of consensus regarding both objectives and procedures. It is difficult to disagree with Jenkinson's (1968) statement that one reason comprehension research lags behind word identification research in productivity lies in the very complexity of the comprehension process itself. Researchers have approached the comprehension process from

at least three directions: from the output, that is, from what the reader can demonstrate of what he knows after he has finished reading; from the reader himself while he is engaged in the act of reading; and from input, or the printed page which initiates the comprehension process. While all of these approaches have yielded some insights into the process itself, the final approach, that which focuses on input, appears to hold the most promise to date, because it deals with the one stable element in an otherwise dynamic process.

Readability Studies - One of the questions which researchers have asked regarding the printed page is: what elements of written language enhance or inhibit reading comprehension? Readability studies have been carried out in order to answer this question in such a way that the answer could be utilized in a formula for measuring the ease or difficulty of written materials. Various researchers have emphasized different written structures as being important. Vogel and Washburne (1928) included the number of prepositions and number of simple sentences as two of their criteria for the Winnetka formula. Gray and Leary (1935) similarly included average sentence length and the number and length of prepositional phrases, as well as the number of first-, second-, and third-person pronouns. The latter factor was discounted by Lorge (1944), although he confirmed the importance of the former two. Dale and Chall (1948) and Spache (1953) emphasized

average sentence length as an important factor of reading ease or difficulty.

The decade of the fifties saw the introduction of the Cloze procedure, which altered the direction of readability studies. The Cloze procedure, as published by Taylor (1953), consists of the systematic deletion of every nth word. Because the syntactic structures remain intact, and are assumed to function as in the normal reading situation, the direct effect of different language elements on reading comprehension can be observed. Consequently, this procedure has come to be the predominant method of measuring readability in reading.

Specific Language Factors Which Could Influence Readability- Without attempting to produce formulas which are characteristic of readability studies, some researchers have focused their attention on specific language factors which could influence reading ease or difficulty. Coleman (1962) examined the possible effect of decreasing sentence length on what he called "comprehensibility". He found that comprehensibility could be improved by raising clause fragments such as subordinate clauses to full sentences or dividing sentences joined by conjunctions (but, for, because, etc.) which signal that the first clause is modified by the second one. However, dividing a sentence which was joined by "and" into two sentences did not improve comprehensibility, and shortening clauses appeared to be more effective in making

them comprehensible than merely emphasizing their boundaries by punctuating them as separate sentences.

Later Coleman (1964) turned his attention to the possible effect of specific grammatical transformations on comprehensibility. In four different experiments, he found that nominalizations, adjectivalizations, and passives were not as comprehensible as their active-verb counterparts. Other researchers have found a significant correlation between passage difficulty and the number of words which occur between a word or phrase and the word or phrase which it modifies (Coleman and Aquino, 1967) and a wide range of linguistic variables, including word depth (Bormuth, 1966).

Specific Language Factors Which Could Influence Comprehension - While the distinction between "comprehensibility" or readability and "comprehension" may at first appear to be an artificial one, it should become clear that the following researchers, while still continuing to focus on the input to the reading process, demonstrate greater concern for and interest in the entire process than did those in the preceding section. While Coleman (1962) questioned whether longer or shorter sentences are easier for adults to comprehend, Robertson (1966) asked what conjunctions had to do with the comprehension which children can demonstrate. She used a modified transformational-

generative grammar to focus on the knowledge of connectives which children in the upper elementary grades possess. She maintained that whether ideas are embedded or conjoined, the association between the ideas is partially contained in the connectives and, consequently, there is a direct relationship between reading achievement and understanding of connectives for that age group.

An analysis of social studies textbooks in an exploratory study by Cossitt (1966) established the predominance of prenominal adjectives in those school materials and led her to conjecture on their contribution to reading comprehension difficulty. This study was followed by a much more intensive one by Fagan (1969) who, utilizing the framework of transformational-generative grammar, investigated the relationship between reading comprehension difficulty and specific language structures found in textbook writing. The purpose of the study was "to investigate the number and types of transformations ... which were found in the written language of three basal reader series at the grade four level, and to determine by means of the 'Cloze' technique, the difficulty which these structures presented for pupils aged nine to twelve in grades four, five, and six (p. iii)." He found that while adjectives were the most common embedded structure in the basal readers, they were among the easiest in terms of reading comprehension difficulty. The two transformations which correlated most

highly with reading difficulty were of the embedding and deletion types. A distinction was made in the study between sentence and passage difficulty, with the conclusion drawn that the presence and difficulty of transformations in sentences had greater effect on their difficulty than did their presence in passages. The difference was accounted for by the greater redundancy found in the passages.

Fagan commented that children "tended to find easier those written structures which one ordinarily found in oral language - vocatives, expletives, direct quotations, and questions." Work by Ruddell (1963) and Tatham (1970) indicated similar findings. Thus it would appear that for the individual to comprehend what he reads he must learn to deal with those syntactic structures of written language which are not ordinarily found in oral language, and that his knowledge is acquired with greater experience in reading.

By focusing on problem readers, Denner (1970) demonstrated that the syntactic competence necessary to integrate and subordinate individual word meaning to sentence meaning is characteristic of average readers at the late grade one level but is missing in problem readers in grade one and grades three to five, as well as in predicted problem readers. The children tested from these groups performed as the average readers from grade one did on tasks associating non-representational linear forms with words, but they could not

synthesize whole sentences from individual linear forms. Denner concluded that the predicted problem readers begin grade one with "an atomistic, mechanistic conception of reading that stresses the relationship of individual graphic forms to the concept of perception and action" and that older problem readers "still fail to subordinate the perceptual-motor meaning of the separate words to the larger linguistic reality of the sentence" (pp. 886-887).

Further evidence that for the mature reader syntactic structures operate on his perception of word sequences was put forward by Forster and Ryder (1971). Their group of forty undergraduate students were rapidly presented word sequences in anomalous, bizarre, and normal semantic conditions. They found that the perception of visually different sentence types was approximately constant for all three semantic conditions, giving additional support to the notion that syntax must be treated differentially in consideration of the reading comprehension process.

When all of the above evidence of the effect of the syntactic structure of written language on reading comprehension is taken into consideration, one must conclude with Little (1972) that "it would seem desirable to develop a means of measuring a child's ability to understand the ways in which structures convey meaning" (p. 20). Early efforts by Gibbon (1941), Strom (1956), and O'Donnell (1962) which focused on the child's ability to see relationships between

the parts of a sentence, knowledge of grammar, and awareness of grammatical structure, respectively, proved disappointing towards this end. In 1970, Bormuth and his associates (1970) commented on the discrepancy between the heavy dependency on transmitting knowledge through written language in the schools and the inadequacies of present techniques to determine whether the children are able to extract meaning from the written language of instructional materials. From this discrepancy has grown a definite need which they state as follows:

> Unfortunately, the testing procedures in current use are unable to provide the information necessary for determining how well students are able to understand the syntactic structures by which language signals meaning. Consequently, there is little knowledge upon which to base the design of instruction for teaching these language-comprehension skills (p. 349).

And it is the instructional needs which Bormuth, Carr, Manning, and Pearson concentrated on in the study which this statement introduced. Working with grade four students, they tested comprehension of specific sentence structures through wh- questions of four different types. Their findings indicated that the structures identified represent homogenous classes of behavior and point toward the possibility that the skills represented by the test tools may be hierarchically related. The students were tested on twenty-five sentence structures which were identified, sixteen intersentence structures and fourteen anaphoric structures.

The authors concluded that their most startling finding was that large proportions of their sixty grade four students were not able to demonstrate comprehension of materials meeting the Dale-Chall criteria of grade four or below through the wh- question forms which are traditionally used in the schools. Their findings strongly support their own argument for an instructional theory of comprehension which takes into account the testing situation which is part of the reality of the classroom.

While Bormuth et al. emphasized the need for an instructional theory, Simons (1970) insisted on the primary need for a sound psychological theory of comprehension. Having concluded from a review of the comprehension literature that research should be based on available linguistic competence theory, he focused on the most important concept of the transformational grammarians, deep structure, and sought to explore the relationship between children's reading comprehension ability and their ability to recover deep structure as indicated by their ability to correctly discriminate between two sentences which are accurate paraphrases of a given lead sentence and one which is not. He found that for the eighty-five grade five students in his sample, scores on his Deep Structure Recovery Test (DSRT) correlated significantly both with scores on the "Cloze" tests he administered and on a standardized reading test which was included in the battery. In fact, scores on the

DSRT accounted for more of the variance in the scores than did I.Q., word knowledge, or word recognition skill.

Both the Bormuth study described above and Simon's study emphasized the need for developing tests which assess the child's knowledge of syntactic structures, and both have brought nearer the possibility of the existence of such tests which are established on objective criteria and which generate responses which can be judged empirically.

Given the large quantity of accrued evidence on the importance of syntactic structures to reading comprehension, it is clear that there is a need for continued work in this area, building toward the day when a cohesive theory of language has been applied and tested in the reading comprehension area and has been translated into classroom teaching and testing practice.

The remainder of this chapter will consist of a discussion of one principle of transformational-generative grammar, structural ambiguity, and those studies which have been related to reading comprehension. These studies provided the framework for the research of Little (1972), from which the present study was generated.

Transformational-Generative Grammar and Ambiguity

In the original form of the theory of transformational-generative grammar, Chomsky (1957) proposed that an adequate grammar must account for the occurrence of certain

phoneme sequences to which two different derivational structures can be assigned. This concept has since been applied to grapheme sequences and, in Little's (1972) words, "the adequacy criterion has been repeatedly stated by transformational grammarians, becoming a canonical example of the superiority of transformational grammar to structuralism or descriptive linguistics" (p. 25). In addition, the ability to perceive ambiguity in a grammatical string has been recognized as one of the basic skills which characterize a mature native speaker of English (Jacobs and Rosenbaum, 1968).

According to MacKay and Bever (1967), transformational grammar defines three levels for the occurrence of ambiguity in sentences--the lexical, the surface structure, and the underlying structure levels. Their descriptions of the three levels are quoted in detail below, as they represent a concept which is central to this study.

> The meanings and sounds of individual words are represented at the lexical level. A sentence is lexically ambiguous if a word or sequence of words has two distinct meanings and no differences at the other grammatical levels. For instance, the sentence The soldiers like the port is lexically ambiguous since the lexical item "port" can mean either "wine" or "harbor".
>
> The manner in which words can be grouped into phrases is represented in the surface structure of sentences. Ambiguity at the surface structure level involves the possibility of two distinct groupings of adjacent words. Consider the sentence Small boys and girls are frightened easily. If the word "small" is grouped with

> "boys and girls" then both the boys and the girls are small. But if "small" is grouped only with "boys" then only the boys are small.
>
> The underlying structure level of sentences represents the essential "logical" relations between words and phrases. For instance, the logical relation between "police" and "drinking" is quite different in these two sentences. The mayor will ask the police to forbid drinking. The mayor will ask the police to cease drinking. Ambiguities at the underlying structure level involve neither a change in meaning of individual words, as in lexical ambiguity, nor a change in the apparent grouping of words, as in surface structure ambiguities, but only a change in the logical relations between words. For example, consider the sentence The mayor will ask the police to stop drinking. In this sentence are the police doing the drinking, or is somebody else? This sentence is ambiguous at the underlying structure level since only the logical relations between police and drinking is altered in the two interpretations (p. 193).

As will be seen below, researchers have claimed to establish the psychological reality of the three levels described by MacKay and Bever. However, the criterion used to differentiate between surface structure ambiguity and underlying structure ambiguity has been brought into question by Prideaux (1972), who has argued that both can be resolved at the surface structure level through labeled bracketing. He claims that:

> The resolution of structural ambiguity does not depend at all on a level of deep structure. It depends only on a notion of semantic representations and paraphrase. Insofar as a notion of autonomous deep syntactic structure (underlying structure) is not required for the resolution of structural ambiguity it is superfluous and unmotivated (p. 10).

It is hoped that this argument among linguists will be resolved in the direction of the greatest economy while preserving the means for evaluating the research which has assumed a difference between surface structure ambiguity and underlying structure ambiguity and, indeed, found differences in the processing of sentences which fit the labels described by MacKay and Bever. It is to this research literature which the present study hopes to contribute; consequently, the traditional labels have been maintained.

Following is a summary of some major pieces of research which have investigated the ability of adults to recognize phonemic or graphemic ambiguity or of identifying the ambiguity of phonemic or graphemic strings.

MacKay and Bever (1967) presented sentences containing lexical, surface structure, underlying structure, or multiple (combinations of these three types) ambiguities to twenty undergraduate students. The sentences, presented on cards, contained eight (plus or minus one) words. The subjects were told that each sentence was ambiguous, and the processing time was measured from the time the sentence was presented until the time the subjects indicated by a "yes" that they had identified both meanings. From the median perception times of each sentence type, a hierarchy of processing time was established. A significant difference

at the .01 level was found for the three types, with lexical ambiguities identified more quickly than surface structure ambiguities and surface structure ambiguities identified more quickly than underlying structure ambiguities.

Even if only one of the ambiguities was identified in multiple ambiguous sentences, they consistently took longer to process.

Another approach taken by MacKay (1966) confirmed the same hierarchical relationship between the four sentence types. Matching unambiguous sentence fragments with ambiguous sentence fragments by making the minimal change necessary to approximate one of the possible meanings, he set up a situation in which undergraduates were asked to orally complete the fragments in complete, grammatical sentences after reading each on a card. The processing time was measured from the time the card was exposed to the subject until he finished saying his completion sentence. The few respondents who reported noticing ambiguity in the fragments during a trial were eliminated from the sample. Consequently, although none of the members of the final sample reported noticing any ambiguities before a trial was completed, they consistently took the least amount of time to complete unambiguous fragments, followed by lexical, derived (i.e., surface) structure, underlying structure, and multiple ambiguities. MacKay argued that this study not only

confirms the finding of the MacKay and Bever (1967) study, but it also complemented it by eliminating the possible effect of differences in structural complexity.

The results of MacKay's experiment were contradicted by the result of a study by Foss, Bever and Silver (1968) which required subjects to determine as quickly as possible whether a picture displayed after a sentence was an instantiation of its meaning. Through pre-tests, they rated each possible meaning of their ambiguous sentences as either "preferred" or "unexpected" according to the number of people who saw that meaning first. In this case, there was no difference between response time to unambiguous sentences and response time to ambiguous sentences if the picture represented the "preferred" meaning of the sentence. If the picture represented the "unexpected" meaning of an ambiguous sentence, the response latency was longer. The authors concluded that under normal circumstances one meaning of the sentence is processed only, with a second processing occurring only if the first is determined to be incorrect. Bever (1969) found support for this order in an experiment which required subjects to paraphrase stimulus sentences as quickly as possible. Ambiguous sentences did not require greater time to paraphrase than did their unambiguous control, when only one meaning was called for.

Carey, Mehler, and Bever (1970) used five unambiguous sentences to establish a set for a particular syntactic structure in other subjects. Then the picture accompanying an ambiguous sentence was presented. The subjects responded most quickly if neither possible meaning was consistent with the picture, more slowly if they claimed to see the ambiguity before responding, and recognized the ambiguity most often if both meanings of the sentence were compatible with the picture. They argued that increasing syntactic expectancy can cause a person to treat an ambiguous sentence as though it were unambiguous.

The apparent contradiction between these investigations into the processing time required by ambiguous sentences was explained by Garrett (1970) as stemming from the nature of the tasks involved. The MacKay and Bever (1967) study tested during the processing of the sentence and showed an effect of ambiguity. Similarly, a study was conducted by Foss (1970) in which response time in a phoneme monitoring task was significantly longer if the phoneme occurred in an ambiguous sentence and the subject was aware of the ambiguity. Another study conducted by Garrett and Lockner and reported by Garrett (1970) showed an effect of ambiguity when an ambiguous sentence monitored by the right ear of the subject was given the interpretation consistent with a disambiguating sentence monitored by the left ear at a 5db lower level. The question of the test task biasing research findings on

ambiguity is one which deserves further investigation.

In his model of speech perception, MacKay (1970) argues in favour of simultaneous processing of possible meanings of ambiguous sentences, although one meaning may be suppressed and the other activated by context. Here again is another issue, how and when ambiguous sentences are processed, which must be resolved by the linguists.

As tentative hypotheses are put forward, however, it is to be hoped that they will be tested on children at various developmental stages, as well as on adults. Even more important, because many of the above studies were conducted by means of oral language, under artificial conditions and with adult subjects, there is a need for additional investigation to determine whether ambiguity in written language, under ordinary conditions, is processed in the same way. The following three studies, the first using oral stimuli and the second two using written stimuli, have investigated the developing awareness of ambiguity in children.

Kessel (1970) included ambiguity among the linguistic constructions for which he tested awareness among children in kindergarten and grades, one, two, three, and five. He selected four sentences for each of the three ambiguity types (i.e., lexical, surface structure, and underlying structure). The sentences were read twice to each child (ten children per grade), so that intonation patterns matched the alternate

meanings. From a set of four drawings, two of which matched the meanings and two of which did not, the child was asked to select the picture or pictures which illustrated the meaning or meanings of the sentence or sentences. He found that his lexical ambiguities were easily interpreted by most of the six-year-olds, and that only the twelve-year-olds showed the same degree of mastery over the structural ambiguities. His subjects showed the greatest gains in detecting underlying structure ambiguity between grades two and three, and the greatest gains in detecting surface structure ambiguity between grades three and five. The limitations of this study were summarized by Little (1972) as follows:

> Kessel recognized a number of limitations of this study, the fact that he used only four sentences of each type; the fact that the sentences used had to be "picturable"; the fact that the difficulty of a particular lexical ambiguity is specific to the lexical item contained therein. To these may be added the fact that no attempt was made to equalize the structural complexity of the sentences across ambiguity types; that one of the twelve sentences contained multiple ambiguity and the pictures suffered from a lack of artistry which may well have complicated the child's comprehension of them (p. 31).

A study by Jurgens (1971), though of a different design, may be considered to be an extension of Kessel's, since it dealt with students in grades seven through eleven. There are difficulties in interpreting her study, however, stemming from a lack of definition of terminology. Her stated intention was "to shed some light on the subject of

the developmental sequence of receptive linguistic competence through a study of the recognition of ambiguity in sentences by students in grades seven, nine, and eleven" (p. 3). "Recognition" is not defined, although it is used to define "perception" of ambiguity: "the ability to recognize two interpretations of an ambiguous sentence when both the ambiguous sentence and the interpretations were presented in written form" (p. 7).

The test which Jurgens constructed to measure the above was composed of sixty lead sentences equally divided among unambiguous, lexical ambiguity, surface ambiguity, and underlying ambiguity, as defined by MacKay and Bever (1967), with two interpretative sentences for each lead sentence. Although the lead sentences were controlled for length (eight words, plus or minus one) and for syntactic complexity, "the analysis of syntactic structures occurring in the sentences did not take into account whether or not the ambiguity was located in the particular structure identified" (Little, p. 32).

Each lead sentence with its accompanying interpretative sentences was typed on a card and presented to each subject who was instructed to read the lead sentence aloud, read the interpretative sentences silently, and then respond with the letter or letters of the sentences which were correct interpretations of the lead sentence. Both the number of correct responses and the time which elapsed from

the first word read of the lead sentence until the response choice was made were recorded. Perception time scores yielded a rank order of difficulty from easiest to most difficult which was: lexical ambiguity, unambiguous, underlying ambiguity, and surface ambiguity. The difference in order of difficulty on Jurgens' test from that found by MacKay and Bever (1967) and MacKay (1966) may stem from differences in the tasks required by the different tests. Correct response scores, on the other hand, when analyzed statistically, revealed no significant differences between the grades for surface ambiguity and significance for underlying structure at the.05 level only when grade seven was compared with both grades nine and eleven.

On the basis of her results, Jurgens concurred with Kessel that the ability to perceive lexical ambiguity cannot be considered on the same developmental level as structural ambiguity, since the former is marked by differences in difficulty of the lexical items themselves. While Kessel found evidence that the ability to perceive underlying structure ambiguity seems to develop first, Jurgens found the ability to perceive underlying ambiguity lagging behind the ability to perceive surface structure ambiguity. This difference undoubtedly lies somewhere between the difference in the ages of the children tested and the difference in the tasks which each test involved. The relationships of perception time and correct response scores to five language

factors were studied, including reading comprehension. Reading comprehension correlated significantly with response time for grades seven and eleven and with correct response scores at grade eleven only, although Jurgens suggested that "these data can only be interpreted with caution" (p. 72).

From the evidence of the above studies, Little (1972) became convinced of the need for more highly controlled investigations into the developmental nature of children's syntactic competency as indicated by their ability to perceive ambiguity in sentences and its relationship to reading comprehension. He states:

> Thus there appears to be some evidence that comprehension of at least structurally ambiguous sentences may shed some light on the comprehension process and seems to reflect a developmental process in children. That this developmental process is tied to an understanding of the grammatical structure of sentences appears evident from the fact that to perceive ambiguity in structurally ambiguous sentences, the individual must be able to perceive two separate syntactic groupings in one grapheme sequence. It thus seems that the structurally ambiguous sentence may well be used as the basis of a measure of the child's linguistic competence as it relates to his reading comprehension of the syntactic structures of written language (pp. 33-34).

Little chose grade five as the most likely place to begin a systematic investigation of children's ability to identify ambiguity in grapheme sequences. Sixty grade five students were administered the Sentence Interpretation Test which he devised. This test contained forty lead sentences. Ten contained surface structure ambiguities of two different

sentence structure types, ten contained underlying structure ambiguities of two different sentence structure types. The consequent four sentence structure types were matched by unambiguous sentences, with twenty unambiguous sentences in all. Each child was given unlimited time to read the lead sentences and indicate whether each of three possible interpretative sentences gave a meaning which was the same as the meaning of the lead sentence or did not. While all twenty of the ambiguous sentences consequently had two interpretative sentences which were paraphrases and one which was not, the interpretative sentences for the unambiguous sentences were varied so that either one, two, or three of the interpretative sentences were paraphrases of the lead sentence. The Reading Comprehension subtest of the Stanford Diagnostic Reading Test, Level II, was administered to each child, yielding subscores for literal and inferential reading comprehension ability.

Although mean scores for the ambiguous sentences were extremely low, a rank order of difficulty was established from least difficult to most difficult as follows: unambiguous, surface structure, and underlying structure. Little concluded the following from his analysis of the data:

> ...grade five students have not generally acquired the ability to identify ambiguity. However, reading comprehension ability was significantly related to identification of ambiguity with the most consistent relationship existing between inferential comprehension ability and the ability to identify surface structure ambiguity (p. iv).

Girls scored consistently higher than boys on the Sentence Interpretation Test.

The groundwork laid by Little calls for the systematic investigation of children at increasing grade levels to determine the rate and pattern of development of the ability to identify structural ambiguity and the relationship this ability has to developing reading comprehension abilities. In addition, the question of the effect of context on the ability to resolve ambiguities in written passages needs to be explored. Halliday (1965) claims that accomplished readers are not aware of ambiguities when they appear in the context of surrounding language, but Menyuk (1971) suggests that context alone is not sufficient. Such an exploration of ambiguities in written context may relate to MacKay's theory that listeners process ambiguities simultaneously, although one possible meaning is suppressed by context. As Little states:

> For children in the acquisition stage of reading, who are not fully aware of the surrounding language, the problems of interpretation that ambiguous sentences create may rely heavily for their solution on syntactic and semantic information. It is the lack of these very factors that characterize the reader in the acquisition stage and thus structural and lexical ambiguities may indeed result in failure to comprehend the meaning of much of the written language to which children are exposed (p. 35).

CHAPTER III

The purpose of this chapter is to describe the design, the sample selected for this study, the experimental and standardized tests used, the pilot study, and the collection and analysis of data.

I. THE EXPERIMENTAL DESIGN

The basic design for the study was a three-way analysis of variance; however, because there were too few subjects per cell, the data were analyzed in terms of three two-way analyses of variance. The independent factors of the model were grade versus reading level, grade versus sex, and reading level versus sex. The dependent variables analyzed were: ability to identify surface structure ambiguity in sentences; ability to identify underlying structure ambiguity in sentences; these two subscores combined as a total ambiguity subscore; ability to identify unambiguous sentences; ability to disambiguate surface structure ambiguity in paragraphs; ability to disambiguate underlying structure ambiguity in paragraphs; total scores on the Sentence Interpretation Test; and total scores on the Contextual Ambiguity Test.

The relationship between Intelligence Quotient (I.Q.) and chronological age (C.A.) and the ambiguity measures were also considered and analyzed through the use of Pearson product-moment correlations.

II. THE SAMPLE

The test population for this study consisted of two grade six classes and two grade seven classes in one school which was assigned to the investigator by the Edmonton Catholic Separate School Board. The total enrollment in the school for grade six was forty-nine; the total enrollment of the two grade seven classes was sixty-two. The school was located in the west end of Edmonton.

It was considered necessary to limit the sample to thirty students (fifteen boys and fifteen girls) in each grade (sixty students in all) who met criteria of language background, reading ability, and intelligence quotient similar to those established by Little (1972), in order for the greatest degree of comparison to be made. Consequently, the following information was gathered in the school by various means:

(a) Language

In order to ensure that results could be compared with Little's (1972) data who worked with native speakers of English, the investigator determined through a questionnaire distributed in the classroom whether the students spoke any language other than English at the time of entering grade one. A copy of this questionnaire is included in Appendix A. Only children who spoke English were eligible for inclusion in the study.

(b) Reading Ability

Little's sample was limited to students in grade five of average reading ability, as determined by the local norms established by the Edmonton Catholic School Board on the school-administered Gates-MacGinitie Reading Test, Survey D (i.e., students whose scores fell within one standard deviation of the district norm for grade fives, 4.1 to 7.5). In order to provide for a greater range of reading abilities, this criterion was changed for the present study. Since the Sentence Interpretation Test was controlled so that the vocabulary difficulty did not exceed grade four level, and since the Contextual Ambiguity Test was controlled so that the vocabulary level did not exceed the grade five level, it was deemed necessary only to eliminate students from the sample if they were not proficient in reading at the grade five level, as determined by the school-administered Gates-MacGinitie Reading Test, Survey D.

(c) Intelligence Quotient

It was also felt necessary to limit the sample by excluding those students whose scores might be affected by low intelligence. For that reason, the Canadian Multi-Level Edition of the Lorge-Thorndike Intelligence Test was administered by the investigator to students who met the language and reading ability criteria. The standard deviation for this test is sixteen I.Q. points; consequently

any student whose I.Q. was recorded as being below 84 was not included in the sample. To make up the sample quota it was necessary to include eight students whose I.Q.'s were recorded as being more than one standard deviation above the mean (i.e., 117 and above). It would not have been possible to fill the sample of grade six students from this one school if these students had been excluded, and it was felt that a wide range of intellectual ability at the upper level would not prejudice the outcome of the study. Table I summarizes the range, means, and standard deviations of the groups included in the final sample.

TABLE I

SUMMARY OF THE CONTINUOUS VARIABLE I.Q.

FOR THE SAMPLE SELECTED

Grade	Mean	Standard Deviation	Range	
			Lowest	Highest
Grade Six	110.77	10.44	90	132
Grade Seven	105.50	9.97	85	126

(d) Grade Level

As Little found that grade five students did not generally have the ability to identify ambiguity in sentences, it was decided to continue the investigation with different students at the grade six and grade seven levels, in the hope of discovering a developmental pattern in the language facilities being tested.

(e) Sex

Because sex differences have been observed as factors influencing language ability and measured by many researchers (Carroll, 1966; Balow, 1963; Weintraub, 1966; Fagan, 1969; Little, 1972), it was considered essential to balance the sample so that there would be an equal number of boys and girls at each grade level.

(f) Chronological Age

Although chronological age was not used as a limiting criterion for selecting the sample, each student's age in months was obtained at the time of the administration of the I.Q. tests. The information recorded at that time is summarized for the groups included in the final sample in Table II.

TABLE II

SUMMARY OF THE CONTINUOUS VARIABLE CHRONOLOGICAL AGE FOR THE SAMPLE SELECTED

Grade	Mean	Standard Deviation	Range	
			Lowest	Highest
Grade Six	140.63	5.02	125	149
Grade Seven	153.43	5.41	146	169

In some cases, there were more than fifteen students (girls in grade six, boys in grade six, girls in grade seven, boys in grade seven) who met the above criteria and completed the tests which were administered by the investigator. In such cases, students were eliminated on a random basis so that thirty students only remained in each grade, divided equally as to sex. The data of these students were used for statistical analysis.

III. TESTING INSTRUMENTS

1. Sentence Interpretation Test

The Sentence Interpretation Test (SIT), constructed by Little (1972) to measure the ability of children to identify the meaning of structurally ambiguous sentences

of written English, was used with slight modification in this study.

Little's test consisted of forty lead sentences: ten with surface structure ambiguity, ten with underlying structure ambiguity, and twenty which were unambiguous. Each lead sentence was followed by three interpretative sentences, one, two, or all three of which gave a meaning of the lead sentence. The subject was asked to read the lead sentence, then read each interpretative sentence in turn and indicate whether it could give a meaning which was the same as the meaning of the lead sentence. A complete description of the test, its construction, its validity and reliability, as it appears in Little's text, can be found in Appendix B.

In an attempt to improve the reliability of the SIT, the investigator chose the seven items from Little's test with the lowest predictability as indicated by the biserial correlations. New items were constructed which matched these items in structure, but which contained different lexical constituents, expanding the test from forty to forty-seven items in total. All subjects were administered this expanded version of the test. On the basis of the item analysis which was performed on the data later, the [redacted] item from the matched pair which appeared to be the least effective predictor was eliminated from the

test data, giving a total once more of forty. Only three of the revised items, all of them unambiguous, actually proved to be better predictors than the originals. The SIT as it was used for data analysis, appropriately renumbered, has been included in Appendix C. A second item analysis was carried out on the forty items retained for the test; this item analysis is summarized in Appendix D. The KR-20 formula reliability coefficient for this administration of the SIT was .841.

2. Contextual Ambiguity Test

The Contextual Ambiguity Test (CAT), used in this study to measure the effect of context on the ability of students to disambiguate structurally ambiguous sentences, was constructed by the investigator. In its final form, it consisted of twenty items, ten for each of the two kinds of structural ambiguities of the SIT. Appendix E of this text contains the CAT in full.

The following criteria were established for the construction of the CAT:

1. Item Selection - The twenty ambiguous sentences from the SIT would be used in their exact form, in order to minimize the effect of the lexical items.

2. Ambiguous Meaning - Each sentence would be embedded in one paragraph which constrained the meaning so that only one of the possible interpretative paraphrases would hold true.

3. Vocabulary - The vocabulary level of the lexical items of the paragraph would be controlled so that only words listed in Carroll's (1971) Word Frequency Book as occurring in reading material of grade five students or below were used.

4. Natural Situation - In order for the paragraphs to simulate naturally-occurring reading situations, the ambiguous sentences in many cases became clauses in longer sentences. However, every effort was made to minimize the use of syntactic structures which, according to Robertson (1966) and Fagan (1970), hinder reading comprehension, while at the same time striving for the most natural-sounding diction. In addition, the position of the ambiguous sentence within the paragraph was varied so that a set for the ambiguous sentence in a certain position would not be established.

5. Paragraph Length - To be sure that unequal length of paragraphs did not bias the person taking the test, it was decided that a consistent paragraph length of sixty words, plus or minus ten, would be maintained. This length was sufficient for providing natural-sounding contextual situations for each ambiguous sentence while at the same time the total length of the test was reasonable for completion within a thirty minute period.

6. Test Format - Following the directions and an example of the test, the items were presented so that the

structurally ambiguous sentences were underlined, within the paragraphs, followed by the same interpretative sentences which occurred in the SIT. Two test items appeared on a page, so that no item would be divided between two pages. The CAT in it entirety is printed in Appendix E.

7. Grammatical and Semantic Acceptability and Effectiveness of Constraints on Ambiguity - All test items were submitted to two groups of people. Fifteen mature native speakers of English were invited to complete the test and make comments as to the grammatical and semantic acceptability of all the paragraphs as well as the effectiveness of each paragraph in constraining the structurally ambiguous sentence contained therein so that only one of the possible meanings would be obvious. In addition, a class of eighteen grade six students was administered the test, and their answers were recorded. The children were invited to comment on those items for which they had found two sentences which gave a meaning for the underlined (structurally ambiguous) sentence or on those items which they found particularly confusing. Items which were universally marked correctly by the adult sample were maintained unchanged. Items which were marked correctly by fewer than ten adults were revised for greater clarity and additional constraint. For items which fell between the parameters of ten to fifteen correct responses from the adult sample, the comments of the adults were considered in the light of the actual performance

of the sample of children, and changes made in cases when they appeared to be justified. The revised items were then re-submitted to a smaller group of the adult sample for final approval.

Instructions for the CAT

The instructions for the test contained one example of a structurally ambiguous sentence which was underlined and embedded in a paragraph which constrained its meaning to only one of the possible meanings. The students were instructed to read each paragraph carefully and then read each of the three interpretative sentences which followed. They were to indicate whether each interpretative sentence gave a meaning for the sentence or part of a sentence which was underlined by placing a check (✓) by each interpretative sentence under a column "GIVES A MEANING" or under a column "DOES NOT GIVE A MEANING".

Validity of the CAT

As this was the initial construction and administration of the test, the only claim for validity that can be made is face validity (Helmstadter, p. 298). Face validity for the CAT as a measure of the ability to disambiguate a structurally ambiguous sentence when it is constrained to only one meaning by the context of a paragraph is claimed on the basis of the following considerations:

1. The ambiguous sentences from the SIT differ structurally from unambiguous sentences (Little, 1972, p. 51).

2. The sentences labeled as containing surface structure ambiguity differ in kind from the sentences labeled as containing underlying structure ambiguity. Little's results, which indicate that children's ability to identify the meanings of sentences with surface structure ambiguity differs from their ability to identify the meanings of sentences with underlying structure ambiguity support the evidence from research by McKay (1966), MacKay and Bever (1967), Kessel (1970), and Jurgens (1971) that individuals perceive the two sentence types differently.
3. It is reasonable to use paraphrases as interpretations of ambiguous sentences since they both have equivalent meanings.
4. The paragraphs in which the ambiguous sentences were embedded were adjudged by a panel of mature college-educated, native speakers of English to effectively constrain the ambiguous sentences so that only one of the possible meanings could apply.
5. The choice by the students of the correct interpretative sentence indicates that they have disambiguated the sentence by recovering one of the possible paraphrases (Simons, 1970; Little, 1972, p. 52).

6. The vocabulary was carefully controlled so that students reading at a grade five level or above could be expected to experience no difficulty with the lexical items. Consequently, the effect of the variables of word recognition which is associated with reading comprehension is minimized to the greatest possible degree.
7. The paragraphs were written in such a way as to simulate natural diction while minimizing the effect of difficult syntactic constructions other than structural ambiguity.
8. The paragraphs were grammatically and semantically acceptable as adjudged by a panel of mature, college-educated, native speakers of English and as was borne out by the pilot study and interviews conducted with certain students in the sample after they had taken the test.

Reliability of the CAT

On the basis of data collected from thirty grade six students and thirty grade seven students, internal reliability as measured by a K-R 20 formula resulted in a coefficient of .768.

3. Stanford Diagnostic Reading Test

Test 1 (Reading Comprehension) of the Stanford Diagnostic Reading Test (SDRT), Level II, Form X was administered to the students to assess their reading

comprehension achievement. The decision was made to use this particular standardized reading test on the following considerations:

1. Since this test was used by Little (1972), the greatest degree of comparability between his study and the present study could be achieved by continuing with the same choice of tests. His rationale for choosing the test is incorporated with the following points.
2. The SDRT has received a favorable review in O.R. Buros The Seventh Mental Measurement Yearbook.
3. Level II of the SDRT is designed for use from the middle of grade four to the middle of grade eight, thus covering the two additional grades, six and seven, included in the present study.
4. The two subskills of literal reading comprehension and inferential reading comprehension as established by the authors of the SDRT provide more accurate analysis of reading comprehension ability than other available standardized tests.
5. The reading comprehension skills are tested on a wide variety of subject-matter materials, including science, social studies, health, etc.
6. The corrected split-half reliability coefficient for the sub-test of reading comprehension at the

grade six level was reported by the author as being .91, and for grade seven as being .92.

7. Content, construct, and concurrent validity are claimed for the test by its authors, largely on the basis of item analysis and correlations between SDRT subtests and the Stanford Achievement Test.

4. Lorge-Thorndike Intelligence Test (Canadian Multi-Level Edition)

The Lorge-Thorndike Intelligence Test (Canadian Multi-Level Edition) was administered to all subjects in the sample in order to obtain the combined verbal and non-verbal numerical Intelligence Quotient (I.Q.) which was used in selecting the sample and for some data analysis. The decision to use this test was based on the following considerations:

1. The editors of this Canadian version of a test which originated in the United States attempted to adjust the cultural bias in order to make it most suitable for English-speaking Canadian children.
2. The various levels of the test are suitable for use in grades three to nine, with levels D and E recommended for use with grades six and seven.
3. The test was normed on a sample of Canadian children who were native speakers of English.

There were 4,000 to 5,000 subjects per grade in the norming sample, which was balanced between the provinces, between Roman Catholic and non-Roman Catholic schools, and by school size in order to be representative of the total population of Canadian English-speaking children.

4. According to the test manual the Odd-Even reliability coefficient for grade six on both the Verbal Battery and Non-Verbal Batter is .911; for grade sevens .872 and .908, respectively.

5. The authors claim construct and concurrent validity for the test on the basis of statistical comparisons between the U.S. versions and other well-known measures of intelligence (e.g. Stanford-Binet and WISC). Data for Canadian students on those other tests are not yet available for carrying out suitable correlations.

As recommended by the test authors, Level D was selected for administration for subjects in grade six, and Level E was selected for administration for subjects in grade seven. The latter choice was made after the school was adjudged to be above the national average as indicated by previous administration of I.Q. tests and standardized reading tests.

IV. PILOT STUDY

A pilot study was conducted in March, 1973, with one class of grade six students in an elementary school within the Edmonton Catholic School System. Its purpose was to gather information concerning the validity of the CAT, as well as to try out the instructions which had been prepared for that test and to determine the amount of time necessary for an average class of grade six students to complete the test.

Eighteen students completed the test at that time, all within a twenty minute period. A discussion with the students revealed that in some cases some had found two sentences which could be considered paraphrases for the underlined sentence in the paragraphs. Some students found some of the paraphrases confusing, and some students reported that the test took a lot of thought. Others reported that they had not found the test difficult.

The data from the pilot study were used for the revision of some test items.

V. COLLECTION OF THE DATA

The investigator obtained information regarding reading ability from the school records. Language background was determined by completion of a questionnaire (see Appendix A) during class time. Chronological age was obtained at the time of administration of the Canadian

Multi-Level Edition of the Lorge-Thorndike Intelligence Test. This test was administered separately over two days to two grade six classes in their regular classrooms. The I.Q. test was administered to a group of thirty-five grade seven students who met the language and reading ability criteria as a group, again over two days. This same group of grade seven students were also given the Stanford Diagnostic Test as a group. Thirty-four grade six students who met the language and reading ability criteria were administered the SDRT as a group.

The Sentence Interpretation Test and Contextual Ambiguity Test were given separately to the four classes, two in grade six and two in grade seven, from which the final sample was drawn. One class in each grade was given the SIT first, and one class was given the CAT first; care was taken to see that the final sample included fifteen students in each grade who had taken the SIT first and fifteen students in each grade who had taken the CAT first.

All tests were administered by the investigator.

VI. DATA ANALYSIS

The data were analyzed according to the following statistical procedures:

1. Two-Way Analysis of Variance

   This procedure was used to test the significance of variation over the five variables associated with the SIT and the three variables associated

with the CAT when the students were divided into six groups. The groupings which were established for this purpose were by sex, by grade, and by classification as high readers or low readers on the basis of the total reading comprehension score.

2. One-Way Analysis of Variance with Repeated Measures

This analysis was used to determine whether there was a significant difference in the variation of the performance of students (when grouped by sex, by grade, or by high or low reading ability according to their total reading comprehension score) on the following variables: surface structure ambiguity, underlying structure ambiguity, total ambiguous sentences, and non-ambiguous sentences from the SIT; surface structure ambiguity and underlying structure ambiguity the CAT; total SIT score and total CAT score.

3. Scheffé Analysis

This analysis was used to determine the source of significant difference in variation where found by the above two analyses.

4. Pearson Product-Moment Correlations

This procedure was applied to determine whether a linear relationship existed:

a. between all variables for the total sample.

b. between all variables for all subjects in grade six.

c. between all variables for all subjects in grade seven.

5. Item Analysis

This analysis was used to determine the item difficulty index and biserial correlation of both the items on the SIT and the items on the CAT.

CHAPTER IV

ANALYSIS OF THE GRADE SIX DATA

After analyzing the data, it became apparent that the grade seven sample performed much lower on the ambiguity subtests of the Sentence Interpretation Test (SIT) than had been anticipated. Therefore an interview was scheduled between six students in grade seven, as well as six students in grade six, and the investigator, on an individual basis. The purpose of the interview was to explore possible factors in the testing situation which would account for the unexpected behavior on the SIT. Student responses were recorded in writing, as it was felt that a taperecorder might inhibit the responses of some of the students.

The interview findings are summarized in Part V of Chapter V. In brief, the responses of the students led the investigator to believe that a large portion of the grade seven subjects had not taken the two language tests (SIT and CAT) seriously. Consequently, the original design which called for analyzing the data as one body with the three dimensions of grade, sex, and reading ability was altered, since it was felt that the results from grade seven would distort the findings of the grade six pupils. It was decided that each grade should be treated individually and any comparisons must be qualified. The order for presenting the data analysis for grade six subjects is as follows:

I. Data Related to Performance of Grade Six Subjects on the Sentence Interpretation Test (SIT).

II. Data Related to Performance of Grade Six Subjects on the Contextual Ambiguity Test (CAT).

III. Comparison of SIT and CAT Data for Grade Six Subjects.

I. DATA RELATED TO THE PERFORMANCE OF GRADE SIX SUBJECTS ON THE SENTENCE INTERPRETATION TEST (SIT).

1. Performance of Grade Six Subjects on the Sentence Interpretation Test (SIT).

Scores on the SIT comprise a measure of student's ability to identify the meaning or meanings of structurally ambiguous sentences and unambiguous sentences of written English by correctly classifying paraphrases of those meanings. The results of grade six subjects for this test are shown in Table III in terms of possible score, mean score, and standard deviations for the total test and subscores of it. The mean scores and standard deviations of Little's grade five sample have been included for comparison.

Total Test. On the total SIT the scores for grade six subjects ranged from 12 to 38 with a mean score of 22.07 and a standard deviation on this score of 6.80. The group mean for grade six subjects was more than four points higher than that of Little's mean for grade five subjects, indicating that the ability the test is designed to measure has developed to a higher degree in grade six students but still has not been adequately acquired.

TABLE III

MEANS AND STANDARD DEVIATIONS FOR THE SIT

FOR GRADE VI

SIT Scores	Possible Score	Mean Score	Standard Deviation	Grade V Mean Score
Total Test	40	22.07	6.80	17.78
Unambiguous Sentences	20	14.97	3.14	11.97
Ambiguous Sentences	20	7.10	4.63	5.85
Surface Structure Ambiguities	10	4.30	2.28	3.48
Underlying Structure Ambiguities	10	2.80	2.71	2.37

Unambiguous Sentences. The scores on those sentences which were unambiguous ranged from 9 to 19 with a mean score of 14.97. This mean score is more than two points higher than the mean score of Little's grade five subjects. However, in both cases, the scores on unambiguous sentences account for approximately 67 percent of the mean scores on the total test, indicating that students from both the grade five and grade six samples could correctly answer those test items containing unambiguous sentences with greater consistency than they could the remainder of the test items which contained structurally ambiguous lead sentences.

Ambiguous Sentences. The scores on those twenty items which contained structurally ambiguous sentences ranged from 2 to 19 with a mean score of 7.10. This mean represents a gain of less than 1.5 points over the mean of the grade five students previously tested, indicating that the ability to identify ambiguity in sentences increases between grades five and six, but only to a small degree.

Surface Structure Ambiguity. Ten of the twenty ambiguous sentences on the SIT contained structures with surface structure ambiguity, and the scores for grade six subjects on these sentences ranged from 1 to 10 with a mean score of 4.30. This mean represents a gain of slightly less than one point over the mean of the grade five sample. However, in both cases, the means on sentences containing surface structure ambiguity account for approximately 60 per-

cent of the mean scores for the total number of ambiguous sentences.

Underlying Structure Ambiguities. The remaining ten ambiguous sentences contained structures with underlying structure ambiguity, and the scores for grade six subjects ranged from 0 to 10 with a mean score of 2.80. This mean represents a gain of slightly less than .5 points over the mean of the grade five sample and indicates the area where least gain occurred.

As can be seen, the data on the SIT for the grade six sample followed a curve which was identical to that of the data for the grade five sample tested by Little, with slight but consistent improvements on all subtests. Consequently, the grade six data support the findings of Little (1972), MacKay (1966), and Jurgens (1971), that structurally ambiguous sentences are more difficult to identify than unambiguous sentences. Furthermore, the fact that sentences with underlying structure ambiguity proved more difficult than sentences with surface structure ambiguity agrees with Little's findings and is in accord with the findings of MacKay (1966), MacKay and Bever (1967), and Kessel's (1970) analysis of correct response scores.

The range of scores for each of the subtests indicates a wide range of linguistic ability within a single grade.

2. Performance of Grade Six Subjects on the Stanford Diagnostic Reading Test (SDRT) - Reading Comprehension Subtest.

Scores on the SDRT reading comprehension subtest were used as a measure of the literal and inferential reading comprehension ability of the students in the sample. Table IV shows the possible score, mean score, and standard deviation for this subtest in total and for both subtests. Means and standard deviations for the grade five sample have been included for ease of comparison.

Total Score. Total scores on the subtest of reading comprehension for the grade six sample ranged from 28 to 52 with a mean score of 42.53. According to the Manual for Administering and Interpreting this test, this mean score represents a grade equivalent of 6.6, whereas the students in the sample were at an actual grade placement of 6.8 at the time of testing. The more than 9 point gain in mean score of the grade six subjects over the mean score of Little's grade five sample is accounted for in part by the difference in sample selection between this study and the previous study. Whereas Little included only subjects who could be classified as average readers because their scores fell within one standard deviation above and below the mean according to the local norms on the Gates-MacGinitie Reading Test, the subjects for the present study had only to meet the criterion of reading at the grade five level or above to be included in the sample, a condition which resulted in a wider range of

TABLE IV

MEANS AND STANDARD DEVIATION FOR THE SDRT

READING COMPREHENSION SUBTESTS FOR

GRADE VI

SDRT Reading Comprehension Subtest	Possible Score	Mean Score	Standard Deviation	Mean For Grade V Sample	Standard Deviation For Grade V Sample
Total Score	60	42.53	8.60	33.32	6.56
Literal Comprehension Score	30	21.70	4.47	18.10	3.63
Inferential Comprehension Score	30	20.17	4.56	15.22	3.83

reading abilities. The standard deviation of 8.60 indicates that, assuming a normal curve, 68 percent of the sample scored between the grade 5.1 and grade 9.6 level as interpreted by the SDRT Manual, and the total range of scores indicates an achievement range from grade 4.3 to grade 10.1.

Literal Comprehension Score. Scores on those thirty items on the SDRT reading comprehension subtest which the authors of this test claim to measure literal comprehension ability ranged from 12 to 28 with a mean score of 21.70. The SDRT manual interprets the extremes of this range as reflecting scores at the first and ninetieth percentile ranks for students placed between the grade 6.5 and grade 7.5 levels.

Inferential Comprehension Score. Scores on those thirty items designed to measure inferential comprehension ranged from 11 to 27 with a mean score of 10.17. The extremes of this range are interpreted by the SDRT manual as reflecting scores at the second and ninetieth percentile ranks for students placed between the grade 6.5 and grade 7.5 levels.

The performance of the grade six students in this sample differed from that of the previous grade five sample in two ways. First, the standard deviations indicate a much wider variation of scores above and below the mean for two-thirds of the students in the sample. Second, the mean scores for both the total test and the two subtests indicate a

higher level of competency in reading comprehension when compared with the students on which the SDRT was normed. According to the Manual for Administering and Interpreting this test, the mean score of the grade five sample for the total test falls at the twenty-second percentile rank for students in grades 5.5 to 6.5, and the mean scores for literal and inferential comprehension fall at the twentieth and twenty-fourth percentile ranks, respectively. For the grade six subjects, however, the mean score for the total test falls at the fortieth percentile rank for students in grades 6.5 to 7.5, and the mean scores for literal and inferential comprehension fall at the thirty-second and forty-second percentile ranks, respectively. These data are summarized in Table V. Both of the differences between the performance of the grade six subjects and the grade five sample described above can be attributed in part to the differences in sample selection, as noted previously. Both groups are similar in being slightly stronger on the subtest of inferential reading comprehension than on the subtest of literal reading comprehension when compared to the norming samples.

TABLE V

SUMMARY OF PERCENTILE RANKING OF MEAN SCORES FOR THE SDRT READING COMPREHENSION SUBTESTS FOR GRADE V AND GRADE VI ACCORDING TO MANUAL FOR ADMINISTERING AND INTERPRETING THE SDRT

SDRT Reading Comprehension Subtest	Grade V Mean Compared to Students in Grades 5.5 to 6.5 in Percentiles	Grade VI Mean Compared to Students in Grades 6.5 to 7.5 in Percentiles
Total Score	22	40
Literal Comprehension Score	20	32
Inferential Comprehension Score	24	42

3. Relationship Between the SIT Variables and Reading Comprehension Variables for Grade Six Subjects.

The Pearson product-moment correlation was used to indicate the probabilities and significance levels of the relationships between each of the continuous variables measured by the SIT and each in turn of the continuous variables measured by the SDRT reading comprehension subtests. As all variables were measured on the interval level and the relationship between the characteristics was assumed to be linear this was deemed to be a suitable method of analysis. These relationships are presented in Table VI.

This analysis revealed that all correlations were positive, all but one reached significance beyond the .01 level, and that one remaining correlation reached significance beyond the .05 level. The highest correlations were between underlying structure ambiguities and total comprehension, between the total SIT scores and total comprehension scores, between ambiguous sentence scores and total comprehension scores, between total SIT scores and inferential comprehension scores, and between total SIT scores and literal comprehension scores. The overall pattern of correlations for the grade six sample differed from that of Little's grade five sample. For the grade six sample, surface structure ambiguities and underlying structure ambiguities correlated at a higher level of significance with all reading comprehension subscores than did unambiguous sentences. This pattern was the reverse of that revealed by analysis of the grade five data.

TABLE

COEFFICIENTS OF CORRELATION BETWEEN THE SIT SCORES AND READING COMPREHENSION SCORES FOR GRADE VI

SIT Scores	Literal Comprehension	Inferential Comprehension	Total Comprehension
Total SIT	.603 **	.611 **	.633 **
Unambiguous Sentences	.394 *	.492 **	.460 **
Ambiguous Sentences	.618 **	.564 **	.618 **
Surface Structure Ambiguities	.554 **	.520 **	.589 **
Underlying Structure Ambiguities	.588 **	.525 **	.[illegible]0 **

** $p < .01$

* $p < .05$

Two factors could contribute to this phenomenon. First, the language ability which is measured by the SIT test and which has developed further between grades five and six may play a more dominant role in reading comprehension in grade six and, second, the broader range of reading abilities which was permitted in the grade six sample may have given a wider range in which the language abilities can make themselves apparent.

Another difference between the data of the two groups was that for the grade six sample (unlike the grade five sample) there were higher correlations between all ambiguity subscores and literal reading comprehension than between these same subscores and inferential reading comprehension. Both grades were similar in that there were higher correlations between unambiguous sentences and inferential reading comprehension than between unambiguous sentences and literal reading comprehension. These differences would appear to qualify Little's conjecture "that instruction in those aspects of language measured by the SIT will help to improve the inferential comprehension ability of grade five students" (p. 71). It may be that instruction in those aspects of language measured by the SIT will improve total reading comprehension abilities, but it is not clear (at least for grade six pupils) that there is as direct a link between inferential reading comprehension and the ability to identify ambiguity in sentences as the grade five data indicated.

4. Differences Between High and Low Readers on the SIT for Grade Six.

In order to further explore the relationship between performance on the SIT and reading comprehension, the total test sample was divided into high and low readers on the basis of their scores on the SDRT Reading Comprehension subtest. The students were divided on the basis of the median score for the total group so that thirty students were designated as high readers and thirty students were designated as low readers. When two two-way analyses of variance were performed, one with reading level as the first factor and, one with grade as the second factor and one with sex as the second factor, it was found that the unexpected grade seven results so strongly affected the interaction that the results from grade six data were obscured. Consequently, the decision was made to treat the data from the two grades separately. Of the thirty subjects in the grade six sample, seventeen fell within the low reader group, and thirteen fell within the high reader group. For the grade seven sample, the numbers in each group were reversed. Although a two-way analysis of variance on data from each grade with high or low reading ability and sex as the two factors would have been desirable in terms of the nature of information generated, it was felt that the small and unequal cell sizes, which would have ranged from 6 to 9, would have seriously limited the reprsentativeness of the resulting interactions. Consequently it was decided

to analyze the differences between high and low readers for grade six subjects on the SIT variables by means of t-tests. This analysis is summarized in Table VII. The analysis revealed a difference between high and low readers in the grade six sample which reached significance beyond the .01 level for all SIT variables. Because the variance for total ambiguous sentences and for underlying structure ambiguities was significant, a Welch prime adjustment of t-tests for unequal variances was used, and both still reached significance beyond the .01 level.

For the grade six sample, high readers consistently scored significantly higher than low readers on the SIT variables. The difference was greatest for the total SIT scores, followed by unambiguous sentences, total ambiguous sentences, and underlying structure ambiguities.

TABLE VII

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN HIGH AND LOW READERS ON SIT VARIABLES FOR GRADE VI

SIT Scores	Group Mean For Low Readers	Group Mean For High Readers	Standard Deviation For Low Readers	Standard Deviation For High Readers	t-Value[1]
Total SIT	18.18	27.15	1.12	3.38	4.581 **
Unambiguous Sentences	12.29	17.15	1.41	2.73	4.073 **
Ambiguous Sentences	4.88	10.00	1.96	5.69	3.106 **
Surface Structure Ambiguities	3.35	5.54	1.41	2.73	2.852 **
Underlying Structure Ambiguities	1.53	4.46	1.12	3.38	3.002 **

** $p < .01$

[1] after Welch prime adjustment of t-tests for unequal variances was used

5. Relationships Among Performance on the SIT, Reading Comprehension Achievement, and Related Variables for Grade Six Subjects.

The contribution of the selected variables, I.Q., sex, and chronological age, to the SIT scores and scores on the SDRT reading comprehension subtest are each discussed below.

a. I.Q. - Little found that for his grade five sample I.Q. was an important factor in the abilities measured by the SIT and the SDRT Reading Comprehension subtests. Correlation coefficients for the grade five sample between I.Q. and the reading comprehension scores were all significant beyond the .01 level. The correlation coefficients for the grade six subjects in the present study between the I.Q. and the reading comprehension scores are reported in Table VIII.

TABLE VIII

COEFFICIENTS OF CORRELATION BETWEEN I.Q. AND READING COMPREHENSION SCORES FOR GRADE VI

Contributing Variable	Literal Comprehension	Inferential Comprehension	Total Comprehension
I.Q.	.418 *	.554 **	.567 **

** $p < .01$
* $p < .05$

For this sample also, I.Q. correlated positively beyond the .01 level of significance with the total comprehension score and with each of the two subscores.

The correlation of I.Q. with the scores on the SIT are reported in Table IX. Once again I.Q. correlated positively beyond the .01 level of significance with the total SIT score and with each of the four subscores. A significantly high, positive relationship was found between I.Q. and underlying structure ambiguities for the grade six sample, while the previous grade five data yielded a positive but nonsignificant relationship between those two variables. The difference may be due to the increase in mean score for the grade six group, particularly for the high readers, for whom the variation in scores is much greater.

The data for these grade six subjects indicate that I.Q. shows a higher relationship to the ability to identify ambiguous sentences than to the ability to identify unambiguous sentences. However, as the difference is slight and the sample is small, this interpretation is made cautiously.

b. Sex - As was noted in Chapter III, many researchers, including Little, have found differences between the sexes on language-related abilities. To establish whether this were true for the grade six subjects of the present study, t-tests on the differences between the mean scores of boys and girls were carried out. The results are summarized in Table X. Although the mean scores for girls were in all cases

TABLE IX

COEFFICIENTS OF CORRELATION BETWEEN I.Q. AND THE SIT SCORES

FOR GRADE VI

Contributing Variable	Total SIT	Unambiguous Sentences	Ambiguous Sentences	Surface Structure Ambiguities	Underlying Structure Ambiguities
I.Q.	.695 **	.570 **	.634 **	.595 **	.581 **

** $p < .01$

TABLE X

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN BOYS AND GIRLS ON SIT VARIABLES FOR GRADE VI

SIT Scores	Group Mean For Boys	Group Mean For Girls	Standard Deviation For Boys	Standard Deviation For Girls	t Value
Total SIT	20.67	23.47	5.74	7.86	1.114
Unambiguous Sentences	14.20	15.73	3.08	3.22	1.334
Ambiguous Sentences	6.47	7.73	3.83	5.51	.731
Surface Structure Ambiguities	3.80	4.80	2.01	2.57	1.188
Underlying Structure Ambiguities	2.67	2.93	2.29	3.24	.260

higher than the mean scores for boys, in no [redacted] was the difference significant. Consequently, it w[redacted]cided that this sample could not contribute conclusiv[redacted]ough evidence to the question of significant differenc[redacted]tween the sexes on language variables to warrant f[redacted]her analysis on the basis of sex.

c. Chronological Age - The students in the grade six sample ranged in age from 125 to 149 months, although the mean age was 140.63 months. The correlation coefficients of chronological age with scores on the SIT and with reading comprehension scores are presented in Tables XI and XII, respectively. All correlations between chronological age and SIT scores were negative and did not reach the .05 level of significance. Between chronological age and reading comprehension scores, one correlation (with literal comprehension) was positive, while the other two were negative, but all were so close to zero as to indicate that no correlation between reading scores and chronological age exists for this sample. These results were very similar to those for the grade five students of Little's study.

It would appear that language development is unique to individuals and that chronological age is not a good predictor of at what stage of development an individual may be. Again results must be interpreted with caution because of the small range within ages.

TABLE XI

COEFFICIENTS OF CORRELATION BETWEEN CHRONOLOGICAL AGE AND THE SIT SCORES FOR GRADE VI

Contributing Variable	Total SIT	Unambiguous Sentences	Ambiguous Sentences	Surface Structure Ambiguities	Underlying Structure Ambiguities
Chronological Age	-.089	-.039	-.105	-.194	-.015

TABLE XII

COEFFICIENTS OF CORRELATION BETWEEN CHRONOLOGICAL AGE AND READING COMPREHENSION SCORES FOR GRADE VI

Contributing Variable	Literal Comprehension	Inferential Comprehension	Total Comprehension
Chronological Age	.026	-.029	-.004

6. Differences Among the Scores on Sentences With Surface Structure Ambiguity, Sentences With Underlying Structure Ambiguity, and Unambiguous Sentences.

The SIT consisted of sentence structures which corresponded to the MacKay and Bever (1967) classifications of surface structure ambiguity and underlying structure ambiguity. To ascertain whether the grade six sample responded in a significantly different way to these two types of structures, as well as to the unambiguous sentences which contained similar, but unambiguous structures, a one-way analysis of variance was carried out using the group means derived from converting each individual's score to a percentage of the possible score. This conversion was done to balance the effect of the greater number of sentences, increasing the likelihood of a higher total score, of the unambiguous types. A summary of this analysis is presented in Table XIII.

TABLE XIII

SUMMARY OF ANALYSIS OF VARIANCE ON SIT SCORES FOR SENTENCES WITH SURFACE STRUCTURE AMBIGUITY, SENTENCES WITH UNDERLYING STRUCTURE AMBIGUITY, AND UNAMBIGUOUS SENTENCES FOR GRADE VI

	Sums of Squares		Mean Squares		df		
Test	Between	Within	Between	Within	Between	Within	F
SIT	3135.144	48050.000	1081.104	800.833	29	60	72.4691 **

** $p < .01$

Since the analysis revealed a difference beyond the .01 level of significance, further analysis was called for to determine where this difference lay. The mean percentage of correct responses for each of the three groups of sentences in the SIT (Table XIV) were compared by means of Scheffé contrasts (Table XV).

TABLE XIV

MEAN PERCENTAGE OF CORRECT RESPONSES FOR SENTENCES WITH SURFACE STRUCTURE AMBIGUITY, SENTENCES WITH UNDERLYING STRUCTURE AMBIGUITY, AND UNAMBIGUOUS SENTENCES FOR GRADE VI

	Surface Structure Ambiguity	Underlying Structure Ambiguity	Unambiguous
Mean Percentage Correct	43.00	28.00	74.83

Since the correlation between surface structure and underlying structure ambiguities reached the .05 level of significance ($r = .41$), it appears that the tasks of surface structure ambiguity and underlying structure ambiguity require different levels of the same skill. Table XIV illustrates once more that the rank order of difficulty, from easiest to most difficult, for this sample was unambiguous sentences, surface structure ambiguity, and underlying structure ambiguity. That these differences all reached significance beyond the .01 level is in complete agreement with the performance of the grade five sample in Little's

TABLE XV

SUMMARY OF SCHEFFÉ CONTRASTS OF MEANS FOR SENTENCES WITH SURFACE STRUCTURE AMBIGUITY, SENTENCES WITH UNDERLYING STRUCTURE AMBIGUITY, AND UNAMBIGUOUS SENTENCES FOR GRADE VI

	Surface Structure Ambiguity 1.	Underlying Structure Ambiguity 2.	Unambiguous 3.
		F	F
1.	--	14.254 **	64.197 **
2.		--	138.9524 **
3.			--

** $p < .01$

study, and it appears that the three tasks of identifying unambiguous sentences, surface structure ambiguity, and underlying structure ambiguity which comprise the SIT require different levels of the same skill.

II. DATA RELATED TO PERFORMANCE OF GRADE SIX SUBJECTS ON THE CONTEXTUAL AMBIGUITY TEST (CAT).

1. Performance of Grade Six Subjects on the CAT.

Scores on the CAT constitute a measure of students' ability to correctly disambiguate the meaning of a structurally ambiguous sentence when that sentence is constrained by the context of a passage so that only one of the possible meanings of the sentence will apply. The results of the grade six subjects for this test are shown in Table XVI.

TABLE XVI

MEANS AND STANDARD DEVIATIONS FOR THE CAT FOR GRADE VI

CAT Score	Possible Score	Mean Score	Standard Deviation
Total Test	20	12.80	3.876
Surface Structure Ambiguities	10	5.73	2.337
Underlying Structure Ambiguities	10	7.07	2.097

Total Test. On the total CAT, the scores for grade six subjects ranged from 4 to 19 with a mean score of 12.80, indicating that for many students in this sample the constraints of context were not sufficient for correctly disambiguating all ambiguous sentences in the test.

Surface Structure Ambiguity. One-half (i.e. ten) of the items on the CAT contained surface structure ambiguities, identical to the ten surface structure ambiguities on the SIT. The scores for grade six subjects on this half of the test ranged from 1 to 9 with a mean score of 5.73. This score accounts only for approximately 45 percent of the group mean for the total test. This is the reverse of the relationship between surface structure ambiguity and the total ambiguous sentences on the SIT and lends support to the notion that the completion of these tests demand different abilities.

Underlying Structure Ambiguity. The remaining ten items contained structures with underlying structure ambiguity which were identical to the ten items containing underlying structure ambiguity on the SIT. The scores for grade six subjects on this half of the test ranged from 2 to 9 with a group mean of 7.07. This mean accounts for approximately 55 percent of the group mean for the total CAT test.

As with the SIT, the wide range of scores for both aspects of the CAT indicates a wide range of linguistic abilities within this grade six sample.

2. Relationship Between the CAT Variables and Reading Comprehension Variables for Grade Six.

Pearson product-moment correlations were used to indicate the probablities and significance levels of the relationships between both of the continuous variables of the GDRT reading comprehension subtest, as discussed in Section 1.2 of this chapter. These relationships are presented in Table [redacted]

This analysis revealed that all correlations were positive, and that two of those involving the total CAT and all of those involving underlying structure ambiguities were significant at least at the .05 level or beyond. The total CAT correlated at the .05 level with literal reading comprehension and beyond the .05 level with the total comprehension scores, the correlation with inferential reading comprehension having failed to reach significance. All correlations between surface structure ambiguities and the reading comprehension variables failed to reach significance, although in this case the correlation coefficient indicating the relationship with inferential reading comprehension is higher than the corresponding correlation coefficient for literal reading comprehension. Underlying structure ambiguities correlated beyond the .01 level of significance with the total comprehension subtest, and beyond the .05 level with the two aspects of that test.

TABLE XVII

COEFFICIENTS OF CORRELATIONS BETWEEN THE CAT SCORES AND READING COMPREHENSION SCORES FOR GRADE VI

CAT Scores	Literal Comprehension	Inferential Comprehension	Total Comprehension
Total Test	.356 *	.349	.425 *
Surface Structure Ambiguities	.171	.192	.266
Underlying Structure Ambiguities	.468 *	.431 *	.489 **

** $p < .01$

* $p < .05$

The correlation coefficient representing the relationship with literal reading comprehension was slightly higher than the corresponding coefficient for inferential reading comprehension, following the same pattern as was observed between underlying structure ambiguity on the SIT and the two reading comprehension subtests.

It is apparent that even when placed in an environment of a constraining context, sentences with underlying structural ambiguities are treated differently by the reader from sentences with surface structure ambiguities.

3. Differences Between High and Low Readers on the CAT for Grade Six.

Just as differences were explored between high and low readers on the SIT variables, so were the CAT variables examined for differences between high and low readers. Table XVIII summarizes the results of t-tests on differences between the two groups on the CAT variables.

The analysis revealed that the difference between the means for the two groups designated as high and low readers was significant beyond the .01 level for the total test and for underlying structure ambiguities. Once more the difference between the groups on surface structure ambiguities failed to reach significance, providing further evidence that the skills required by the two kinds of sentences when constrained by context are different in some way. Furthermore,

TABLE XVIII

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN HIGH AND LOW READERS ON CAT VARIABLES FOR GRADE VI

CAT Scores	Group Mean For Low Readers	Group Mean For High Readers	Standard Deviation For Low Readers	Standard Deviation For High Readers	t-Value
Total Test	11.29	14.77	3.85	3.22	2.624 **
Surface Structure Ambiguities	5.18	6.46	2.27	2.40	1.498
Underlying Structure Ambiguities	6.12	8.31	1.90	1.80	3.201 **

** $p < .01$

the abilitiy to disambiguate underlying structure ambiguities when constrained by a single context appears to be more discriminating of reading achievement at this level.

4. Relationships Among Performance of the CAT and Related Variables.

a. I.Q. - The correlation coefficients between I.Q. and the variables for the CAT are reported in Table XIX.

TABLE XIX

COEFFICIENTS OF CORRELATION BETWEEN I.Q. AND THE CAT SCORES FOR GRADE VI

Contributing Variable	Total CAT	Surface Structure Ambiguities	Underlying Structure Ambiguities
I.Q.	.458 **	.354	.453 *

** $p < .01$

* $p < .05$

For the grade six sample, there was a positive correlation between I.Q. scores and all of the CAT variables, reaching significance beyond the .01 level with the total test and significance beyond the .05 level with underlying structure ambiguities. Once more a difference between surface structure ambiguity and underlying structure ambiguity and their relationship to I.Q. scores is apparent.

b. Sex - To determine whether there were significant differences between the performance of boys and girls on the variables of the CAT, t-tests were carried out on the means of the two groups. These are summarized in Table XX. Although the girls obtained higher scores than the boys on all aspects of the CAT, the differences failed to reach significance.

TABLE XX

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN BOYS AND GIRLS ON CAT VARIABLES FOR GRADE VI

CAT Scores	Group Mean For Boys	Group Mean For Girls	Standard Deviation For Boys	Standard Deviation For Girls	t-Value
Total Test	11.67	13.93	3.81	3.86	1.618
Surface Structure Ambiguities	5.27	6.20	2.40	2.34	1.078
Underlying Structure Ambiguities	6.40	7.73	2.06	2.05	1.775

c. Chronological Age--The coefficients of correlation between the variable chronological age and the various aspects of the CAT are presented in Table XXI. For the grade six sample, very low positive correlations were found between chronological age and the various aspects of the CAT. Once more age would not appear to be a good predictor of the development of linguistic ability which is measured by the CAT for such a limited sample.

TABLE XXI

COEFFICIENTS OF CORRELATION BETW[illegible] CHRONOLOGICAL AND THE CAT SCORES FOR GRADE VI

Contributing Variable	Total CAT	Surface Structure Ambiguities	Underlying Structure Ambiguities
Chronological Age	.109	.048	.148

5. Differences Among the Scores on Passages With Surface Structure Ambiguity and Passages With Underlying Structure Ambiguity.

To determine whether the difference between scores on passages containing surface structure ambiguity and passages containing underlying structure ambiguity was greater than could have occurred by chance, a one-way analysis of variance was carried out on the means after the individual scores had been converted to percentages of the total possible scores. The results of this analysis are reported in Table XXII. The analysis revealed that the differences between the scores on the two kinds of structurally ambiguous sentences which were embedded in paragraphs in the CAT, surface structure ambiguity and underlying structure ambiguity, were significant beyond the .01 level. This is further evidence that the two kinds of ambiguous structures require different levels of the same reading skill even when they are constrained by the context of a larger written passage.

TABLE XXII

SUMMARY OF ANALYSIS OF VARIANCE ON CAT SCORES FOR PASSAGES WITH SURFACE STRUCTURE AMBIGUITY, AND PASSAGES WITH UNDERLYING STRUCTURE AMBIGUITY FOR GRADE VI

	Sums of Squares		Mean Squares		df		
Test	Between	Within	Between	Within	Between	Within	F
CAT	22540.000	9700.000	777.241	323.333	29	30	10.995 **

** $p < .01$

III. COMPARISON OF SIT AND CAT DATA FOR THE GRADE SIX SUBJECTS

The two experimental tests used in this study, the SIT and the CAT, were designed to measure two different, though related, language skills. A student who could correctly identify paraphrases of the ambiguous sentences on the SIT indicated that he was able to perceive that two different groupings of adjacent words, in the case of surface structure ambiguities, or two logical groupings of words, in the case of underlying structure ambiguities, would yield two different meanings. To correctly select a paraphrase of an ambiguous sentence on the CAT, however, he must perceive that only one of the possible meanings of the sentence could hold true within the constraints of the contextual paragraph. The latter task is much more common in the experience of students with their daily reading experiences, and was assumed by the investigator to be less difficult than the skill required for children to perform on the SIT. To determine whether there were significant differences between the scores on the corresponding ambiguity subscores of the SIT and the total CAT scores for the grade six students, t-tests were performed on the mean scores for that group. The results of this analysis are summarized in Table XXIII.

TABLE XXIII

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN SCORES ON SPECIFIC SIT VARIABLES AND CAT VARIABLES FOR GRADE VI

Variables Compared	Mean For SIT Variable	Mean For CAT Variable	Standard Deviation For SIT Variable	Standard Deviation For CAT Variable	t-Value
Total Ambiguous Sentences (SIT) With Total (CAT)	7.10	12.80	4.63	3.88	6.697 **
Surface Structure Ambiguities (SIT With CAT)	4.30	5.73	2.28	2.34	3.174 **
Underlying Structure Ambiguities (SIT With CAT)	2.80	7.07	2.71	2.00	7.807 **

** $p < .01$

The differences between scores on the total ambiguous sentences of the SIT and the total CAT and for each of the corresponding subscores were significant beyond the .01 level. The rank order of difficulty between surface structure ambiguity and underlying structure ambiguity was reversed on the two tests, with scores for the former being higher on the SIT and lower on the CAT, and the t-values for the difference between the scores on those portions of the two tests containing surface structure ambiguities was much less than for the difference between those portions of the tests containing underlying structure ambiguities. Thus it appears that the presence of a larger context makes sentences containing underlying structures easier in the sense that children score higher, this, however, is not the case with sentences containing surface structure ambiguity.

To determine whether scores on the ambiguous portions of the SIT would be good predictors of the ability to disambiguate sentences within a contextual framework (as may be determined by the degree of relationship of scores), Pearson product-moment correlations was carried out on the means for the grade six group on the corresponding subscores of the two tests. The results of this analysis, and the significant levels, are reported in Table XXIV.

TABLE XXIV

COEFFICIENTS OF CORRELATION BETWEEN THE SIT SCORES AND CAT SCORES FOR GRADE VI

SIT Scores	Total CAT	Surface Structure Ambiguities (CAT)	Underlying Structures Ambiguities (CAT)
Total Ambiguous Sentences (SIT)	.430 *		
Surface Structure Ambiguities (SIT)		.309	
Underlying Structure Ambiguities (SIT)			.365 *

* $p < .05$

Low but significant, positive relationships were found to exist between those portions of the two tests containing underlying structure ambiguities and between the total ambiguity scores on the SIT and the total CAT scores.

This analysis reinforces the suggestion that the tasks of identifying surface structure ambiguity may require different competencies than the tasks of disambiguating structurally ambiguous sentences containing surface structure ambiguity. The tasks of dealing with underlying structure ambiguity, whether on the SIT or CAT, appear to be dependent on a common factor. Since the mean scores for surface structure ambiguity on the SIT and CAT were much closer than the underlying structure ambiguity scores for those tests, it appears that the use of context is less effective in helping grade six students disambiguate surface structure ambiguity than underlying structure ambiguity. This may be partly due to the cues (the grouping of words) for signalling ambiguity in the surface-type structures which may be more predominant and not as readily influenced by context as the logical relationships existing between words in the underlying structures.

IV. SUMMARY OF FINDINGS

The findings resulting from the interpretation of the test data are summarized as follows:

1. Scores on the SIT indicate that grade six students are better able to identify the paraphrases of

unambiguous and ambiguous sentences than grade five students.

2. The order of difficulty for the sentence types on the SIT was the same for grade six as it was for grade five, and was as follows from easiest to most difficult: unambiguous sentences, surface structure ambiguity, and underlying structure ambiguity.

3. Scores on the CAT indicate that grade six students are better able to disambiguate underlying structure ambiguities when these occur in paragraphs.

4. There was a significant, positive relationship between reading comprehension as measured by the SDRT and the language abilities measured by both the SIT and CAT for grade six students.

5. High readers in grade six performed significantly higher on all aspects of the SIT and on the underlying ambiguity portion of the CAT than did low readers.

6. I.Q. correlated significantly with reading comprehension scores and with the SIT for grade six, as it did for the previous grade five sample. I.Q. also correlated significantly with the CAT scores.

7. Girls in grade six scored higher than boys on all variables of both the SIT and CAT. These differences, however, did not reach the level of significance set for this study.

8. Chronological age did not correlate significantly with the SIT scores, the CAT scores, or with the reading comprehension scores for the grade six subjects.

9. The difference among scores for unambiguous sentences, surface structure ambiguities, and underlying ambiguities on the SIT and for surface structure ambiguities and underlying structure ambiguities on the CAT were significant for students in grade six, thus lending further support to the results of those studies (MacKay, 1966; MacKay and Bever, 1967; Kessel, 1970; Jurgens, 1971; Little, 1972) which indicates similar differences between these different groups of structures.

10. Grade six students found the tasks of disambiguating surface structure and underlying ambiguities in context significantly less difficult than the tasks of identifying the two meanings possible in isolated sentences.

11. There was a low but significant correlation for grade six students between scores on the total SIT and CAT tests and between scores on those portions of the tests containing underlying structure ambiguities.

CHAPTER V

ANALYSIS OF THE GRADE SEVEN DATA; COMPARISON OF GRADE SIX AND GRADE SEVEN ON THE SIT AND CAT; INTERVIEW DATA

The analyses of data found in this chapter will be presented in the following order:

I. Data Related to Performance of Grade Seven Subjects on the Sentence Interpretation Test (SIT).

II. Data Related to Performance of Grade Seven Subjects on the Contextual Ambiguity Test (CAT).

III. Comparison of SIT and CAT Data for Grade Seven Subjects.

IV. Comparison of Data for Grades Six and Seven on the SIT and CAT.

V. Interview Data.

I. DATA RELATED TO PERFORMANCE OF GRADE SEVEN SUBJECTS ON THE SIT.

All analytical procedures which were performed on the data for the SIT for grade six subjects were also performed on the grade seven data for the SIT. Consequently, this section will consist of the systematic reporting of those analyses, following the format established in Section I of Chapter IV. Explanations of tests and procedures which were explained there will not be repeated in this section.

1. Performance of Grade Seven Subjects on the SIT.

The results of the grade seven subjects on the SIT are shown in Table XXV in terms of possible score, mean score, and standard deviations for the total test and for all its subscores. The means for the grade six subjects and for the grade five sample have been included for ease of comparison.

Total Test. On the total SIT the scores for grade seven subjects ranged from 9 to 34 with a mean score of 21.00 This mean was slightly more than three points higher than the mean score for Little's grade five sample and nearly one point less than the mean score for the grade six subjects previously discussed. Furthermore, the scores for grade seven subjects were more tightly clustered around the mean than were the scores for the grade six subjects, with approximately 68 percent of the scores falling between the scores of 15.93 and 26.07 while 68 percent of the grade six scores fell between the scores of 15.27 and 28.87.

TABLE XXV

MEANS AND STANDARD DEVIATIONS FOR THE SIT FOR GRADE VII

SIT Scores	Possible Score	Mean Score	Standard Deviation	Mean For Grade VI	Mean For Grade V
Total Test	40	21.00	5.07	22.07	17.78
Unambiguous Sentences	20	15.83	3.29	14.97	11.97
Ambiguous Sentences	20	5.17	3.03	7.10	5.85
Surface Structure Ambiguities	10	3.77	1.87	4.30	3.48
Underlying Structure Ambiguities	10	1.40	1.50	2.80	2.37

The skills tested by the SIT would appear to be more highly developed in the grade seven subjects than in the grade five sample, but not as highly developed as in the grade six sample.

Unambiguous Sentences. The scores on those twenty sentences which were unambiguous ranged from 7 to 20 with a mean score of 15.83. This mean score is exactly three points higher than the mean score for the grade five sample and nearly one point higher than the score for the grade six subjects. While unambiguous sentences accounted for approximately 67 percent of the total score for both of the latter groups, they accounted for approximately 75 percent of the total score for the grade seven subjects. Clearly this group were able to correctly answer the test items containing unambiguous sentences with greater consistency than they could the remainder of the test items which contained structurally ambiguous lead sentences.

Ambiguous Sentences. The scores on the twenty items which contained structurally ambiguous sentences ranged from 0 to 16 with a mean score of 5.17, nearly .50 points less than the mean score for the grade five sample and nearly two points less than the grade six subjects for total ambiguous sentences.

Surface Structure Ambiguity. The scores for grade seven subjects on the ten items which contained surface structure ambiguity ranged from 0 to 8 with a mean score of

3.77. This score is .30 points higher than the mean score of the grade five sample and nearly .50 points less than the grade six subjects. Furthermore, while scores on items containing surface structure ambiguity accounted for approximately 60 percent of the total ambiguity scores for both of the latter groups, this score accounts for approximately 73 percent of the total ambiguity score for the grade seven subjects.

Underlying Structure Ambiguity. For the remaining ten sentences which contained underlying structure ambiguity, the scores for the grade seven subjects ranged from 0 to 8 with a mean score of 1.40. This score was nearly 1 point less than the mean score for the grade five sample and nearly 1.5 points less than the mean score for the grade six subjects.

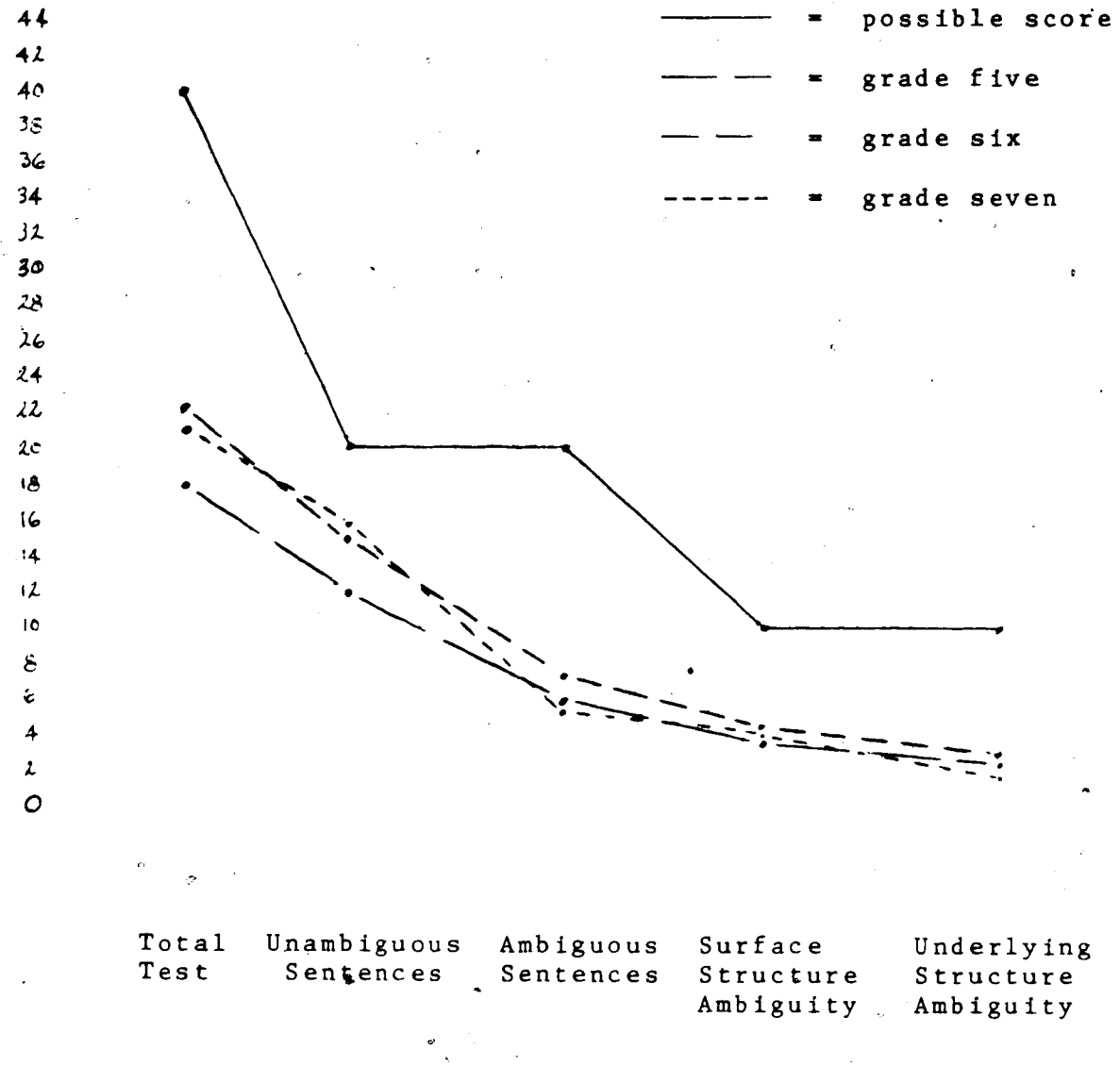
It is clear that this grade seven sample found underlying structure ambiguity to be the most difficult type of item on the test, although their total test performance did not follow the pattern established by the previous groups. Figure 1 illustrates the relationships between the scores on the test variables for the three grades. The information in this figure raises the question of why the grade seven subjects behaved in an unexpected manner on the SIT. Possible answers to this question will be discussed in Section V of this chapter.

FIGURE 1

SUMMARY OF SCORES FOR GRADE V,

GRADE VI, AND GRADE VII

ON SIT



2. Performance of Grade Seven Subjects on the Stanford Diagnostic Reading Test (SDRT) —Reading Comprehension Subtest.

The means and standard deviations for the grade seven subjects on the SDRT reading comprehension subtests are presented in Table XXVI.

Total Test - Scores on the total test for grade seven subjects ranged from 31 to 57 with a mean score of 46.03. According to the Manual for Administering and Interpreting this test, this mean score represents a grade equivalent of 7.5, whereas the students in the sample were at an actual grade placement of 7.8 at the time of testing. Since the grade equivalent of the mean for the grade six subjects was 6.6, it would appear that the difference between the average reading comprehension ability of the groups was not equal to one year. However, the standard deviations were different, 5.79 for grade seven subjects compared to 8.60 for grade six subjects, indicating that there was a wider range of reading comprehension abilities among the grade six subjects. The grade seven subjects were more tightly clustered around the mean, 68 percent having scores which would place them between grades 6.0 and 9.7 in grade equivalents. The total range of scores represented an achievement range for grade six subjects of between 4.8 and 12.3.

TABLE XXVI

MEANS AND STANDARD DEVIATIONS FOR THE SDRT READING COMPREHENSION SUBTEST FOR GRADE VII

SDRT Reading Comprehension Subtest	Possible Score	Mean Score	Standard Deviation	Mean For Grade VI	Mean For Grade V
Total Score	60	46.03	5.79	42.53	33.32
Literal Comprehension Score	30	23.60	2.90	21.70	18.10
Inferential Comprehension Score	30	22.47	3.40	20.17	15.22

Literal Comprehension Score. Scores on those thirty items in the SDRT reading comprehension subtest which the authors of this test claim to measure literal comprehension ability ranged from 18 to 29 with a mean score of 23.60. The SDRT manual interprets the extremes of this range as reflecting scores at the eighth and ninety-fourth percentile ranks for students placed between the 7.5 and 8.5 grade levels.

Inferential Comprehension Score. Scores on the thirty items designed to measure inferential comprehension ranged from 13 to 28 for grade seven subjects, with a mean score of 22.47. The extremes of the range are interpreted by the SDRT manual as reflecting scores at the first and ninety-fourth percentile ranks for students between the grade 7.5 and grade 8.5 levels.

According to the Manual for Interpreting and Administering the SDRT, the performance of the grade seven subjects when compared to the population on which the test was normed was similar to the performance of the grade six subjects when compared to the norming population. The mean scores for the grade seven subjects on the inferential comprehension subtest ranked at approximately the forty-sixth percentile when compared with students in grade 7.5 to 8.5; the mean score for literal reading comprehension ranked at approximately the thirtieth percentile when compared to the same norming group.

These percentile ranks are only slightly different from those of the grade six subjects when compared to students in grades 6.5 to 7.5. These relationships are presented in Table XXVII. It would appear that the reading comprehension abilities of the two groups, grade six and grade seven, of the sample were similar.

TABLE XXVII

SUMMARY OF PERCENTILE RANKING OF MEAN SCORES FOR THE SDRT READING COMPREHENSION SUBTESTS FOR GRADE VI AND GRADE VII ACCORDING TO MANUAL FOR ADMINISTERING AND INTERPRETING THE SDRT

SDRT Reading Comprehension Subtest	Grade VI Mean Compared to Students in Grades 6.5 to 7.5 in Percentiles	Grade VII Mean Compared to Students in Grades 7.5 to 8.5 in Percentiles
Total Score	40	40
Literal Comprehension Score	32	30
Inferential Comprehension Score	42	46

3. Relationship Between the SIT Variables and Reading Comprehension Variables for Grade Seven Subjects.

Table XXVIII presents the coefficients of correlation between all the SIT variables and the reading comprehension variables. All correlations were close to zero, although some were negative and some were positive. Since these results were very different from the results for both the grade five sample and the grade six subjects, and since the performance of the grade seven subjects on the SDRT followed an expected pattern, it is the performance on the SIT of those same subjects which appear to be atypical.

TABLE XXVIII

COEFFICIENTS OF CORRELATION BETWEEN THE SIT SCORES AND READING COMPREHENSION SCORES FOR GRADE VII

SIT Scores	Literal Comprehension	Inferential Comprehension	Total Comprehension
Total SIT	-0.09	.05	-0.02
Unambiguous Sentences	-0.04	-0.03	-0.04
Ambiguous Sentences	-0.10	0.10	0.01
Surface Structure Ambiguities	-0.24	0.08	-0.08
Underlying Structure Ambiguities	0.08	0.12	0.11

4. Differences Between High and Low Readers on the SIT for Grade Seven.

The analysis of the differences between high and low readers for grade seven subjects on the SIT variables by means of t-tests is summarized in Table XXIX. In all cases, the low readers did better as a group than the high readers, although the larger standard deviations for the latter group indicate a wider range of scores for 68 percent of those subjects. In no case did the difference reach a level of significance. It would appear that the unexpected results from the grade seven subjects on the SIT are due in large part to the atypical performance of the high reader group among the grade seven subjects.

TABLE XXIX

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN HIGH AND LOW READERS ON SIT VARIABLES FOR GRADE VII

SIT Scores	Group Mean For Low Readers	Group Mean For High Readers	Standard Deviation For Low Readers	Standard Deviation For High Readers	t-Value
Total SIT	22.31	20.00	3.84	5.89	-1.224
Unambiguous Sentences	16.54	15.29	2.26	3.96	-1.011
Ambiguous Sentences	5.77	4.71	2.42	3.57	-0.933
Surface Structure Ambiguities	4.31	3.35	1.80	1.93	-1.381
Underlying Structure Ambiguities	1.46	1.35	0.97	1.87	-0.190

5. Relationships Among Performance on the SIT, Reading Comprehension Achievement, and Related Variables For Grade Seven Subjects.

a. I.Q. - The relationships between the variable I.Q. and the variables of the reading comprehension subtest and the SIT test are presented in Table XXX and Table XXXI, respectively. As with the grade five sample and the grade six subjects, I.Q. correlated beyond the .01 level of significance with all variables of the reading comprehension subtest. However, unlike the previous two groups, I.Q. correlated at a .05 level of significance with only one of the SIT variables, unambiguous sentences. The other correlations were not significant, and one, with surface structure ambiguities, was negative, although it was so close to zero as to indicate no correlation at all.

TABLE XXX

COEFFICIENTS OF CORRELATION BETWEEN I.Q. AND READING COMPREHENSION SCORES FOR GRADE VII

Contributing Variable	Literal Comprehension	Inferential Comprehension	Total Comprehension
I.Q.	.519 **	.591 **	.617 **

** $p < .01$

TABLE XXXI

COEFFICIENTS OF CORRELATION BETWEEN I.Q. AND THE SIT SCORES FOR GRADE VII

Contributing Variable	Total SIT	Unambiguous Sentences	Ambiguous Sentences	Surface Structure Ambiguities	Underlying Structure Ambiguities
I.Q.	.312	.379 *	.111	-.013	.241

* $p < .05$

b. Sex - The results of t-tests which were performed on the differences between mean scores for boys and girls in grade seven on the SIT variables are summarized in Table XXXII. For all variables but one, unambiguous sentences, the mean scores for girls were higher than the mean scores for boys. However, in no cases were the differences significant.

TABLE XXXII

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN BOYS AND GIRLS ON SIT VARIABLES FOR GRADE VII

SIT Scores	Group Mean For Boys	Group Mean For Girls	Standard Deviation For Boys	Standard Deviation For Girls	t-Value
Total SIT	20.93	21.07	.68	2.06	-0.070
Unambiguous Sentences	16.33	15.33	1.18	2.42	0.814
Ambiguous Sentences	4.60	5.73	1.50	4.10	-1.006
Surface Structure Ambiguities	3.40	4.13	2.47	4.06	-1.056
Underlying Structure Ambiguities	1.20	1.60	3.24	6.68	-0.714

c. Chronological Age - The relationships between chronological age and the variables of the reading comprehension subtests and the SIT are presented in Table XXXIII and Table XXXIV, respectively. The coefficients of correlation between chronological age and all variables on both tests were negative; however, all failed to reach a level of significance. This is in keeping with the findings for the grade six subjects.

TABLE XXXIII

COEFFICIENTS OF CORRELATION BETWEEN CHRONOLOGICAL AGE AND READING COMPREHENSION SCORES FOR GRADE VII

Contributing Variable	Literal Comprehension	Inferential Comprehension	Total Comprehension
Chronological Age	-.203	-.05	-.136

TABLE XXXIV

COEFFICIENTS OF CORRELATION BETWEEN CHRONOLOGICAL AGE AND THE SIT SCORES FOR GRADE VII

Contributing Variable	Total SIT	Unambiguous Sentences	Ambiguous Sentences	Surface Structure Ambiguities	Underlying Structure Ambiguities
Chronological Age	-.243	-.251	-.134	-.105	-.141

6. Differences Among the Scores on Sentences With Surface Structure Ambiguity, Sentences With Underlying Structure Ambiguity, And Unambiguous Sentences.

A summary of the analysis of variance between the scores for the three kinds of sentence structures contained in the SIT for the grade seven subjects is presented in Table XXXV.

TABLE XXXV

SUMMARY OF ANALYSIS OF VARIANCE ON SIT SCORES FOR SENTENCES WITH SURFACE STRUCTURE AMBIGUITY, SENTENCES WITH UNDERLYING STRUCTURE AMBIGUITY, AND UNAMBIGUOUS SENTENCES FOR GRADE VII

	Sums of Squares		Mean Squares		df		
Test	Between	Within	Between	Within	Between	Within	F
SIT	14768.062	75883.375	509.243	1264.723	29	60	178.746 **

** $p < .01$

Since this analysis revealed a difference beyond the .01 level of significance, further analysis was called for to determine where this difference lay. The percentage of correct responses for each of the three groups of sentences in the SIT (Table XXXVI) were compared by means of Scheffé contrasts (Table XXXVII).

TABLE XXXVI

MEAN PERCENTAGE OF CORRECT RESPONSES FOR SENTENCES WITH SURFACE STRUCTURE AMBIGUITY, SENTENCES WITH UNDERLYING STRUCTURE AMBIGUITY, AND UNAMBIGUOUS SENTENCES FOR GRADE VII

	Surface Structure Ambiguity	Underlying Structure Ambiguity	Unambiguous
Mean Percentage Correct	37.67	14.00	79.17

TABLE XXXVII

SUMMARY OF SCHEFFÉ CONTRASTS OF MEANS FOR SENTENCES WITH SURFACE STRUCTURE AMBIGUITY, SENTENCES WITH UNDERLYING SENTENCE AMBIGUITY, AND UNAMBIGUOUS SENTENCES FOR GRADE VII

	Surface Structure Ambiguity 1.	Underlying Structure Ambiguity 2.	Unambiguous 3.
1.	--	46.001 **	144.451 **
2.		--	348.780 **
3.			--

** $p < .01$

The rank order of difficulty for the grade seven subjects was the same as for the grade six subjects and for the grade five subjects in Little's study. The contrasts of the means, which all reached a level of significance beyond the .01 level, indicate that even though the grade seven subjects did not perform as expected, their efforts lend credence to the supposition that the three different kinds of sentence structures which were included in the test called for different levels of the same skill for successful completion.

II. DATA RELATED TO PERFORMANCE OF GRADE SEVEN SUBJECTS ON THE CONTEXTUAL AMBIGUITY TEST (CAT)

All analytical procedures which were performed on the data for the CAT for grade six subjects were also performed on the grade seven data for the CAT. Consequently, this section will consist of the systematic reporting of those analyses, following the format established in Section II of chapter IV. Explanations of tests and procedures which were explained there will not be repeated in this section.

1. Performance of Grade Seven Subjects on the CAT.

The results of the grade seven subjects on the CAT are shown in Table XXXVIII.

Total Test. On the total CAT the scores for grade seven subjects ranged from 5 to 19 with a mean score of 12.50. This score was .30 points less than the mean score for the grade six subjects.

Surface Structure Ambiguity. For the ten items which contained surface structure ambiguities, the scores for the grade seven subjects ranged from 1 to 9 with a mean score of 4.97. This score is nearly one point less than the mean for the grade six subjects. While scores for this portion of the test represented approximately 45 percent of the group mean for the total test for the grade six subjects, it represents only 39 percent of the group mean for the grade seven subjects.

TABLE XXXVIII

MEANS AND STANDARD DEVIATIONS FOR THE CAT FOR GRADE VII

CAT Scores	Possible Score	Mean Score	Standard Deviation	Grade VI Mean Score
Total Test	20	12.50	3.95	12.80
Surface Structure Ambiguities	10	4.97	2.32	5.73
Underlying Structure Ambiguities	10	7.53	1.89	7.07

Underlying Structure Ambiguity. On the remaining ten items which contained structures with underlying structure ambiguity, the scores for grade seven subjects ranged from 4 to 10 with a mean score of 7.53. This is nearly .50 points higher than the mean for grade six subjects for this portion of the test and represents 61 percent of the group mean for the total test. While the differences in scores between the two groups was slight, the order of difficulty for the two kinds of ambiguous structures when constrained by the context of a paragraph remained the same: surface structure ambiguity was more difficult than underlying structure ambiguity.

2. Relationship Between the CAT Variables and Reading Comprehension Variables for the SDRT.

The relationships between both of the continuous variables measured by the CAT and each in turn of the continuous variables of the SDRT reading comprehension subtests are reported in Table XXXIX. Although some correlations were positive and some were negative, all were so close to zero as to indicate no correlation at all, with the possible exception of the one between underlying structure ambiguities and literal comprehension. This coefficient, however, also failed to reach the level of significance set for the study.

TABLE XXXIX

COEFFICIENTS OF CORRELATION BETWEEN THE CAT SCORES AND READING COMPREHENSION SCORES FOR GRADE VII

CAT Scores	Literal Comprehension	Inferential Comprehension	Total Comprehension
Total Test	.032	-.017	.011
Surface Structure Ambiguities	-.032	-.078	-.060
Underlying Structure Ambiguities	.105	.060	.096

3. Differences Between High and Low Readers on the CAT for Grade Seven.

Table XI summarizes the results of t-tests on the differences between scores of high readers and low readers on the CAT variables. In no case did the difference between high and low readers on a variable of the CAT reach a level of significance. The group means for high and low readers were very close, although high readers as a group performed higher on every variable.

TABLE XL

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN HIGH AND LOW READERS ON CAT VARIABLES FOR GRADE VII

CAT Scores	Group Mean For Low Readers	Group Mean For High Readers	Standard Deviation For Low Readers	Standard Deviation For High Readers	t-Value
Total Test	11.77	13.06	3.47	4.42	.866
Surface Structure Ambiguities	4.62	5.24	2.10	2.58	.708
Underlying Structure Ambiguities	7.15	7.82	1.68	2.10	.942

4. Relationships Among Performance on the CAT and Related Variables.

a. I.Q. - The correlation coefficients between I.Q. and the variables for the CAT are reported in Table XLI.

TABLE XLI

COEFFICIENTS OF CORRELATION BETWEEN I.Q. AND THE CAT SCORES FOR GRADE VII

Contributing Variable	Total CAT	Surface Structure Ambiguities	Underlying Structure Ambiguities
I.Q.	.113	.070	.150

All correlations were positive, but none reached a level of significance. These findings are consistent with the performance of the grade seven subjects on the ambiguous portions of the SIT but not consistent with the performance of the grade six subjects on the CAT.

b. Sex - The results of the t-tests which were performed on the differences between mean scores for boys and girls in grade seven on the CAT variables are given in Table XLII. The mean scores for girls were higher on all variables than the mean scores for boys, but in no case were the differences significant.

TABLE XLII

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN BOYS AND GIRLS ON CAT VARIABLES FOR GRADE VII

CAT Scores	Group Means For Boys	Group Means For Girls	Standard Deviations For Boys	Standard Deviations For Girls	t-Value
Total Test	11.67	13.33	4.12	3.89	1.140
Surface Structure Ambiguities	4.60	5.33	2.29	2.44	.848
Underlying Structure Ambiguities	7.07	8.00	1.94	1.85	1.34

c. Chronological Age - The coefficients of correlation between the variable chronological age and the various aspects of the CAT are presented in Table XLIII.

TABLE XLIII

COEFFICIENTS OF CORRELATION BETWEEN CHRONOLOGICAL AGE AND THE CAT SCORES FOR GRADE VII

Contributing Variable	Total CAT	Surface Structure Ambiguities	Underlying Structure Ambiguities
Chronological Age	-.284	-.305	-.221

All correlations were negative, and none reached a level of significance. It would appear that within a particular grade level, chronological age is not a good predictor of the level of ability attained by students on those language skills measured by either the SIT or the CAT. However, the same restrictions of a small range of scores as occurred for the grade six students also holds for the grade seven students.

5. Differences Among the Scores on Passages With Surface Structure Ambiguity and Passages With Underlying Structure Ambiguity.

The results of the one-way analysis of variance which was carried out on the scores for grade seven subjects are reported in Table XLIV. The analysis reveals that the grade seven scores for the CAT, which follow the pattern of the grade six scores for that test, were also significantly different on the two types of structures involved in the test. Because the behavior of this group on the language tests appears to have been atypical, no conclusion can be drawn from this observation. A discussion of the results of the follow-up interview which was conducted to explore the atypical behavior of the grade seven group on the language tests follows in Section V.

TABLE XLIV

SUMMARY OF ANALYSIS OF VARIANCE FOR PASSAGES WITH SURFACE STRUCTURE AMBIGUITY AND PASSAGES WITH UNDERLYING STRUCTURE AMBIGUITY FOR GRADE VII

	Sums of Squares		Mean Squares		df		
Test	Between	Within	Between	Within	Between	Within	F
CAT	23475.00	23475.00	809.482	441.667	29	30	85.076 **

** p < .01

TABLE XLV

SUMMARY OF t-TESTS ON DIFFERENCES BETWEEN SCORES ON SPECIFIC SIT VARIABLES AND CAT VARIABLES FOR GRADE VII

Variables Compared	Mean For SIT Variable	Mean For CAT Variable	Standard Deviation For SIT Variable	Standard Deviation For CAT Variable	t-Value
Total Ambiguous Sentences (SIT) With Total CAT	5.17	12.50	3.03	3.96	9.184 **
Surface Structure Ambiguities (SIT With CAT)	3.77	4.97	1.87	2.31	2.320 *
Underlying Structure Ambiguities (SIT With CAT)	1.40	7.53	1.50	1.89	16.983 **

** $p < .01$

* $p < .05$

III. COMPARISON OF SIT AND CAT DATA FOR THE GRADE VII SUBJECTS.

To ascertain whether there were significant differences between the scores for the grade seven subjects on the corresponding ambiguity subtests of the SIT and CAT and between the total ambiguity scores for the SIT and the total CAT scores, t-tests were performed on the mean scores of that group. The results of this analysis are summarized in Table XLV. The differences between the mean scores reached significance beyond the .01 level for the total SIT and total CAT and for those portions of the tests containing underlying structure ambiguities. The difference between the means on those portions of the tests containing surface structure ambiguities reached significance beyond the .05 level. Although the latter significance level is lower than for the grade six subjects, the pattern of a lower t-value for the differences on those portions of the SIT and CAT containing surface structure ambiguities remains the same.

The results of the Pearson product-moment correlation which was carried out on the scores for the grade seven sample on the corresponding subtests of the two tests are reported in Table XLVI. All correlations were low and positive, with one, between those portions of the SIT and CAT containing underlying structure ambiguities, reaching significance beyond the .05 level. Even though the grade seven group performed differently from the grade six group

TABLE XLVI

COEFFICIENTS OF CORRELATION BETWEEN THE SIT SCORES AND CAT SCORES FOR GRADE VII

SIT Scores	Total CAT	Surface Structure Ambiguities (CAT)	Underlying Structure Ambiguities (CAT)
Total Ambiguous Sentences (SIT)	.265		
Surface Structure Ambiguities (SIT)		.129	
Underlying Structure Ambiguities (SIT)			.360 *

* $p < .05$

and their overall performance is not considered to be typical of their language ability, this pattern of correlations is similar to that of the grade six group. Again the inference can be drawn that the tasks of either identifying or disambiguating underlying structure ambiguities are more similar than the corresponding tasks for surface structure ambiguities. It would appear also that, as for the grade six students, surface structure ambiguities may present a greater potential barrier to reading comprehension for the lower reading achiever.

IV. COMPARISON OF DATA FOR GRADES SIX AND SEVEN ON THE SIT AND CAT.

To determine whether there were significant differences between the scores for the grade six subjects and the grade seven subjects on the SIT and the CAT, one-way analyses of variance were conducted. The results of these analyses are reported in Table XLVII and Table XLVIII, respectively. The grade six subjects performed significantly better on that portion of the SIT containing underlying structure ambiguities than did the grade seven subjects. Their scores were sufficiently higher than the grade seven group to cause the difference between their scores for the total ambiguous sentences on the SIT and those of the grade seven group to reach a level of significance, as well. The differences between the two groups for all of the variables of the CAT failed to reach a level

TABLE XLVII

SUMMARY OF ANALYSIS OF VARIANCE ON DIFFERENCES BETWEEN GRADE VI AND GRADE VII FOR THE SIT VARIABLES

SIT	Sums of Squares	Mean Squares	df	F
Total Test	33.664	33.664	1	1.236
Unambiguous Sentences	7.674	7.074	1	2.834
Ambiguous Sentences	71.544	71.544	1	5.583 *
Surface Structure Ambiguities	5.579	5.579	1	.238
Underlying Structure Ambiguities	37.165	37.165	1	.267 **

** $p < .01$

* $p < .05$

TABLE XLVIII

SUMMARY OF ANALYSIS OF VARIANCE ON DIFFERENCES BETWEEN GRADE VI AND GRADE VII FOR THE CAT VARIABLES

CAT	Sums Of Squares	Mean Squares	df	F
Total Test	5.620	5.620	1	.383
Surface Structure Ambiguities	11.767	11.767	1	2.153
Underlying Structure Ambiguities	1.122	1.122	1	.305

of significance.

The following section contains a discussion of data collected during the follow-up interviews conducted to explore possible causes for the atypical behavior of the grade seven subjects on the SIT and CAT tests.

V. INTERVIEW DATA

As was discussed at the beginning of Chapter IV, twelve students were selected for individual interviews with the investigator in an attempt to explore the possibilities of why the scores of the grade seven subjects on the language tests, SIT and CAT, were different from what had been expected. The section of the chapter provides the criteria used for the selection of the students and a discussion of the content and findings of the interviews conducted.

Selection of the Students for Interview

Since the object of the interview was to explore possible causes for unexpectedly low scores on the ambiguous sentences of the SIT and discrepant scores on the CAT, students were chosen on the basis of their scores on the tests according to the following pattern:

1. Three students in grade seven who were in the bottom half of the total group on the Stanford Diagnostic Reading Test (SDRT) Reading Comprehension scores and who had

the lowest scores on the Total Ambiguity subtest of the SIT of all the grade seven students whose reading scores also fell in the bottom half of the total group.

2. Three students in grade seven whose scores on the SDRT fell in the top half of the scores of the total group for reading comprehension and whose scores on the Total Ambiguity subtest of the SIT were the lowest of all the grade seven students whose reading comprehension scores were among the top half of the total group.

3. Three students in grade six whose scores on the SDRT fell on the bottom half of the scores of the total group for reading comprehension and whose scores on the Total Ambiguity subtest of the SIT were the lowest of all the grade six students whose reading comprehension scores were among the bottom half of the total group.

4. Three students in grade six whose scores on the SDRT fell in the top half of the scores of the total group for reading comprehension and whose scores on the Total Ambiguity subtest of the SIT were the lowest of all grade six students whose reading comprehension scores were among the top half of the total group.

Content and Results of the Interview

The interviews were constructed so as to acquire additional information within four areas of interest.

1. Readministration of SIT and CAT. In order to determine whether the answers of these students were representative of their normal language behavior or were peculiar to the testing situation, the investigator selected certain items from the SIT and CAT for readministration. In each case, two items containing surface structure ambiguity, two items containing underlying structure ambiguity, and two items containing unambiguous sentences from the SIT which were missed by that student (by identifying only one of the possible meanings on the case of the ambiguous sentences) in the original test administration and which were missed by the greatest number of students as indicated by the item difficulty index were selected for readministration.

Items were similarly selected from the CAT so that two items containing surface structure ambiguity and two items containing underlying structure ambiguity which were missed by that student in the original test administration and which were missed by the greatest number of students as indicated by the item difficulty index were readministered. The items selected from the CAT and SIT for readministration for each student are shown in Appendix G. Two of the grade six students (I and L in Appendix G) did not retake two items from the underlying ambiguity portion of the CAT, because they had obtained scores of 9 and 10 (10 being the maximum possible) during the original administration of the test.

Both were in the high reader category for grade six students but were among the lowest scores on the SIT for that group. One student (I) was readministered the only passage containing underlying ambiguity which he had missed, and the other (K) did not retake any items from the CAT containing underlying structure ambiguity.

The directions for each test were reviewed by the investigator, and the items selected were administered within a reasonable length of time.

The results for all twelve students for both tests are summarized in Figure 2. On the six items from the SIT, the range for the grade seven students was 0 - 3 points, with a mean score of 1.50. For the grade six students, the range for these items was 0 to 3 points, with a mean score of 1.34. Thus it may be seen that the grade seven students interviewed did slightly better as a group than the grade six students. On the four items from the CAT, the grade seven students scored from 0 to 4 with a mean score of 1.50. For the grade six students on this test, scores ranged from 0 to 3 with a mean score of .83. Again the grade seven students did better as a group than the grade six students interviewed by retaking selected items from the CAT. This finding is not surprising, since the total grade seven sample performed much lower than was expected on the original administration of the tests. However, it should

FIGURE 2

ADDITIONAL CORRECT RESPONSES BY CATEGORY

Student	Reader Category	Grade	SIT				CAT		
			US	SS	UA	TAS	US	SS	TC
A	HR	7	0	1	2	3	2	2	4
B	HR	7	0	1	2	3	0	0	0
C	LR	7	0	0	2	2	0	0	0
D	LR	7	1	0	0	1	1	1	2
E	LR	7	0	0	0	0	0	1	1
F	HR	7	0	0	0	0	1	1	2
G	LR	6	0	1	1	2	0	0	0
H	HR	6	0	0	0	0	1	0	1
I	HR	6	1	1	1	3	0	0	0
J	LR	6	0	0	0	0	1	2	3
K	HR	6	0	1	1	2	-	0	0
L	LR	6	0	0	1	1	0	1	1

US = Underlying Structure
SS = Surface Structure
UA = Unambiguous Sentences
TAS = Total Ambiguous Sentence
TC = Total CAT

be noted that most of the gain for grade seven students on the SIT is accounted for by items from the unambiguous sentence category. Only one grade seven student and one grade six student marked one additional item from the underlying structure ambiguity category correct. Two grade seven students marked one additional item from the surface structure ambiguity category correct, while three grade six students did likewise.

When the same scores are examined for differences between high and low readers in the interview subjects, there is no apparent difference on the CAT, since both groups have a mean score for the four items of 1.17. However, on the SIT, the high readers scored higher than the low readers, with a mean score of 1.84 compared to a mean score of 1.00 for the latter group. Furthermore, it is the high readers in grade seven who contributed most to this group score, indicating that the unusual scores for that group in the original testing situation were probably not representative of the actual language ability required for coping with the skills which the SIT measured.

2. Explanation of Answers. Following completion of the two test retakes, each student was given an opportunity to explain why he had marked the sentences as either giving a meaning or not giving a meaning which was the same as the lead sentence, in the case of the SIT, or as the underlined sentence or part of a sentence in the case of the

CAT. In cases of the subject's having changed his answer from what it was on the original test, he was asked if he recalled what his original answer had been and, if he did, why he had answered differently.

From the explanations of the students' answers, it was clear that most of the students read the ambiguous sentences in the SIT only one way; that is, they read the lead sentence, perceived a meaning, and looked for a paraphrase which fit that meaning. This was particularly true of the grade seven students. They felt that the second paraphrase which described the alternative meaning of the sentence could not mean the same as the lead sentence because it could not fit the first paraphrase meaning which they had selected. It is interesting that many changed their answers on the retake from their original incorrect response of only one of the two possible meanings to the second possible meaning, this time omitting the paraphrase which they had originally perceived as being true. It thus appears that at any point in time a particular meaning tends to predominate and this meaning strongly influences how the subject perceives the relatedness of structures conveying similar semantic information.

Some students also displayed difficulties with the paraphrases. Their tactic was to look for the same words in the interpretative sentences, rather than focusing on the overall meaning of the sentences. Consequently, they indicated

that "Cats and little dogs like to go exploring" could mean the same as "Little cats and dogs like to go exploring". This difficulty with paraphrase was especially apparent with unambiguous sentences. More than one student argued that the sentence "That the island was discovered pleased the king" could not mean the same as "The discovery of the island pleased the king" because of the different suffixes on the word "discover". Unawareness on the part of students of the flexibility of English as a language in allowing different ways of saying things and unawareness of the functions of variant forms of words suggest a specific need for language instruction, especially because of the implications of loss of reading comprehension when unfamilar forms are used.

Although they were specifically instructed not to, a few students based their answers on their experiences with the real world. Thus, one student claimed that "Salesmen like pleasing most people who shop or they would not be salesmen".

3. Recognition of Ambiguity. When the student's second attempt at an ambiguous sentence was also wrong in that he had selected only one of the possible paraphrases or had selected one paraphrase and the distractor statement which was not a paraphrase of the ambiguous sentence, the interviewer used [illegible] several approaches to determine why he was having [illegible] and whether he could recognize

ambiguity. a) He was asked to read the sentences aloud to determine whether he was reading the words as they were written. b) The investigator read to him an identical sentence with different lexical items, and he was asked to tell what the meaning was. If he gave only one meaning, he was asked if there were any other possible meanings. c) The investigator provided a context in which the other meaning would apply, and the student was asked if the sentence could give a meaning which would hold true in that context.

a. Most of the students were asked to read aloud the ambiguous sentence and the interpretative sentences. It became clear that many of them misread at least one of the sentences involved. The miscues made in reading accounted for some students' marking the interpretative sentence "Only those ladies who were like growing flowers came" as meaning the same as "Only those ladies who liked growing flowers came" since they read both as having identical words (omitting "who" and changing "like" to "liked" in the first sentence). Reading aloud also caused some students to discover that a lead sentence such as "Michael's brother watched the game in the patio" and the distractor sentence, "While on the patio Michael watched his brother's game", did not mean the same because, although they contained the same lexical items, the word order gave them different meanings.

b. In some cases, when students heard one ambiguous sentence which was identical in structure to the second but with different lexical items, they were able to supply two different meanings. Most, however, required the investigator's repeating the sentence with different intonation patterns before they could supply two meanings.

c. Only one student (E in Figure 2) insisted that alternative contexts provided by the examiner were not sufficient to cause the other correct paraphrase to hold true. However, all other students when given an alternative context for items that had been incorrect, not only were able to find the correct answers for these contexts but also immediately generalized their strategies to the original items which they could then deal with successfully without any discussion on the part of the examiner. It appeared to the investigator that these students had "learned" the concept of ambiguity during the experience and would have been able to apply it to another test if they had been given the opportunity.

4. The Testing Situation. The investigator discussed the testing situation with each student and asked what conditions could have been changed to improve his performance on the original administration of the tests.

The grade six subjects who were interviewed did not have much to suggest as to ways in which the original testing situation could have changed so that they would have scored

ambiguity. a) He was asked to read the sentences aloud to determine whether he was reading the words as they were written. b) The investigator read to him an identical sentence with different lexical items, and he was asked to tell what the meaning was. If he gave only one meaning, he was asked if there were any other possible meanings. c) The investigator provided a context in which the other meaning would apply, and the student was asked if the sentence could give a meaning which would hold true in that context.

a. Most of the students were asked to read aloud the ambiguous sentence and the interpretative sentences. It became clear that many of them misread at least one of the sentences involved. The miscues made in reading accounted for some students' marking the interpretative sentence "Only those ladies who were like growing flowers came" as meaning the same as "Only those ladies who liked growing flowers came" since they read both as having identical words (omitting "who" and changing "like" to "liked" in the first sentence). Reading aloud also caused some students to discover that a lead sentence such as "Michael's brother watched the game in the patio" and the distractor sentence, "While on the patio Michael watched his brother's game", did not mean the same because, although they contained the same lexical items, the word order gave them different meanings.

higher. However, the grade seven students were very talkative on the subject. Student E admitted that he had not taken seriously the original testing on the SIT and the CAT and he did not believe most of his friends had. Furthermore he had not taken the retest seriously, either. All grade seven subjects interviewed said that since the tests were not to be counted for their marks and would not have any influence on their teacher (as the I.Q. test and the standardized reading test presumably could have), they had treated them with indifference, searching only for the easiest "right" answer with the least amount of effort. Some felt that an element of competition between the two grade seven classes or between the grade seven classes and the grade six classes would have challenged them more and caused them to be more thoughtful as they completed the test items. These responses caused the investigator to conclude that the grade seven scores as a whole were not representative of the ability of those students for the language skills which the tests were designed to measure. It is in this sense that the grade seven sample were considered atypical.

VI. SUMMARY OF FINDINGS

The findings resulting from the interpretation of the test data are summarized as follows:

1. Scores on the SIT indicate that grade seven students are able to identify the paraphrases of unambiguous

sentences more easily than either grade five or grade six students and ambiguous structures not as well as grade six students but better than grade five students.

2. The order of difficulty for the sentence types on the SIT remained constant for grades five, six and seven and was, from easiest to most difficult: unambiguous, surface structure ambiguity, underlying structure ambiguity.

3. Scores on the CAT indicate that grade seven students, like their grade six counterparts, are better able to disambiguate underlying structure ambiguities than surface structure ambiguities when they occur in constraining paragraphs.

4. There were no significant relationships between reading comprehension as measured by the SDRT and the abilities measured by the SIT and CAT for grade seven students.

5. Low readers in grade seven scored higher than high readers on the SIT and high readers scored higher on the CAT, but the differences were not significant. The interview data support the assumption that the scores for high readers in grade seven were atypical of their language abilities tested by the SIT or the CAT.

6. I.Q. correlated significantly with reading comprehension ability, but not with the SIT and CAT scores, for the grade seven subjects.

7. Girls in grade seven scored higher than boys on all but one subscore of the SIT and on each of the CAT variables, although none of these differences reached the level of significance set for this study.

8. Chronological age did not correlate significantly with the SIT scores, the CAT scores, or with the reading comprehension scores for the grade seven subjects. These results were similar to those of the grade six analyses and were not unexpected because of the small range of ages within each grade.

9. Similar to the grade six results, grade seven students differed significantly on their scores on unambiguous, surface structure, and underlying structure portions of the SIT and on surface structure and underlying structure portions of the CAT.

10. Grade seven students were better able to cope with ambiguity when the structures were embedded in paragraphs than when they appeared in isolated sentences. This was similar to the performance of the grade six group.

11. Scores for the grade seven students on the portions of the SIT and CAT containing underlying structure ambiguities correlated at a level of significance set for this study.

12. Grade seven students were better able to disambiguate underlying structure ambiguities than grade six students and grade six students scored higher than grade seven students in disambiguating surface structure ambiguities on the CAT. The interview data suggest that the scores of grade seven students on the CAT did not accurately reflect their ability with the language skill which was being tested.

13. Grade six students scored significantly higher than grade seven students on the underlying ambiguity and the total ambiguity portions of the SIT. No other scores were significantly different.

CHAPTER VI

SUMMARY, CONCLUSIONS AND IMPLICATIONS

A review of the purpose and deign of the study is presented in this chapter, as well as a summary of the findings and conclusions from the analyses. Implications for education and suggestions for further research are also presented.

I. SUMMARY

One purpose of this study was to examine the performance of grade six and grade seven students on two language tests, one of which had been administered previously to a group of grade five students, for evidence of developmental acquisition of the linguistic competence which the tests were designed to measure. A second purpose was to determine whether a relationship existed between the language abilities measured (the ability to identify structurally ambiguous sentences and the ability to disambiguate those same sentences when they are read within the constraints of a contextual paragraph) and reading comprehension ability.

To achieve this purpose, the Sentence Interpretation Test (SIT) which was devised for the previous study with grade five students by Little (1972) was revised by the investigator. This test was designed to measure the individual's ability to identify the meanings of ambiguous and

unambiguous sentences by means of paraphrases of those sentences. A second language test, the Contextual Ambiguity Test (CAT), was constructed by the investigator, and provided a paragraph for each of the ambiguous sentences occurring in the SIT so that only one of the possible meanings held true. The Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest, was used to measure reading comprehension ability.

The sample selected consisted of thirty grade six students and thirty grade seven students who spoke only English at the time they began their schooling, were proficient in reading at the grade five level or above, and were of average or above in I.Q. Testing was conducted in April, 1973. The investigator conducted all tests and scored them personally.

Pearson product-moment correlations were used to determine relationships between the scores on the tests and the relationships between the abilities measured by the tests and the contributing variables of I.Q. and chronological age. Analyses of variance were used to determine differences for students within each grade between boys and girls and between high and low readers on the abilities of identifying ambiguity in isolated sentences and disambiguating ambiguous sentences embedded in constraining paragraphs, as well as for differences between the various sentence structures found in the two language tests and differences between the performance of

students in grade six and students in grade seven on the two language tests.

A follow-up interview was conducted with six students in each grade to explore possible reasons for the atypical behavior of the grade seven group on the two language tests.

II. FINDINGS AND CONCLUSIONS ON TEST DATA

The null hypotheses found in Chapter I are given below and data for and against the statement of each hypothesis is discussed.

Null Hypotheses

Null Hypothesis I

There is no significant difference between the scores of grade six students and grade seven students on

a) Sentence Interpretation Test (SIT)

b) Contextual Ambiguity Test (CAT)

The one-way analysis of variance conducted on the scores for the grade six and grade seven students revealed a difference which reached the level if significance set for rejection of the null hypotheses in this study for two subscores of the SIT (underlying structure ambiguities and and total ambiguous sentences), in favor of the grade six subjects. Thus this portion of the null hypothesis is rejected. Scores for grade seven students were higher than grade six students on the total SIT test, the unambiguous

sentences of the SIT, and on that portion of the CAT containing underlying structure ambiguities. Differences between these scores, however, were not significant.

The information gathered during the interview which followed initial analysis of the data suggested that no conclusions regarding the development of the linguistic abilities measured by the SIT and CAT can be drawn from these findings, since the grade seven group apparently did not take the two language tests seriously. Consequently, it cannot be assumed that their scores reflected the level of ability which they had achieved in the language skills measured by the SIT and CAT.

Scores for the grade six subjects were higher than scores for Little's grade five subjects on all subscores of the SIT, including the score for the total test, indicating that there is a development between grades five and six in the skills measured by the SIT.

Null Hypothesis II

There is no significant difference between the scores of students in both grade six and grade seven on the Contextual Ambiguity Test and the ambiguous portion of the Sentence Interpretation Test.

The t-tests which were performed on the differences between ambiguous portions of the SIT and the corresponding portions of the CAT revealed differences at least at the

.05 level. The null hypothesis is rejected. Results indicate that both grade six and grade seven students found the task of disambiguating ambiguous sentences when constrained by the context of a paragraph to be significantly easier than the task of identifying the two paraphrases of ambiguous sentences in isolation. However, because neither group scored in a way which would reflect total mastery of the skill required to disambiguate sentences in context, it is not clear whether context alone is sufficient for a reader to successfully deal with ambiguous structures. The passages were written to constrain the meaning of the ambiguous sentence to a single interpretation and this factor alone may have accounted for the differences observed. Or it could be that context in general provides for more redundancy in terms of information necessary for comprehension. Fagan (1969) gave the latter as a possible explanation for the findings of his study which showed that subjects could better comprehend transformational structures in larger passages as opposed to sentences in isolation.

Null Hypothesis III

There is no significant relationship between scores on a test of reading comprehension (Stanford Diagnostic Reading Test, Level II, Reading Comprehension subtest) and scores on a test of identification of ambiguity (Sentence Interpretation Test) or scores

on a test on the ability to disambiguate structurally ambiguous sentences in context (Contextual Ambiguity Test) for students in grade six and for students in grade seven.

The correlations between reading comprehension ability and the scores attained by the grade seven students on the SIT and CAT failed to reach a level of significance and were close to zero. The null hypothesis cannot be rejected for grade seven. It is not possible to draw a conclusion from this finding, however, since the performance of the grade seven students appears to have been atypical in light of the information gained in the interview. For the grade six subjects, on the other hand, correlations between reading comprehsnsion ability and scores on the SIT and CAT were significant and positive, thus resulting in the rejection of null hypothesis for grade six. These findings followed the trend set by the grade five students in Little's (1972) study except that for the grade five group inferential reading comprehension correlated at a higher level of significance with surface structure ambiguities on the SIT than for the grade six group where inferential reading comprehension correlated higher with underlying structure ambiguities. Thus while it is apparent that the ability of grade six (and grade five) subjects to cope with ambiguous sentences is related to their reading comprehension ability, the specific nature of this relationship across grade levels is not clear. It may

be that identifying surface structure and underlying structure ambiguities represent two different skills for students in grade six. The analysis did not reveal a significant relationship between reading comprehension abilities and the ability to disambiguate structurally ambiguous sentences in context for the grade six students.

Null Hypothesis IV

There is no significant difference between high and low readers in grade six and grade seven on their ability to

a) identify ambiguity

b) disambiguate structurally ambiguous sentences in context

The t-tests performed on the group mean scores for grade six subjects revealed that high readers found all tasks on the SIT and the task of disambiguating underlying structure ambiguities on the CAT easier than low readers. Differences reached the .01 level and consequently allow the rejection of this part of the null hypothesis. High readers in grade seven did not score as high as low readers on the SIT but performed higher on the CAT. The differences between high and low readers in grade seven did not reach the level of significance set for this study. This part of the null hypothesis cannot be rejected.

The results of this hypothesis tend to confirm the results of the previous one with regard to the relationship

between the language abilities measured by the SIT and CAT and reading achievement. There appears to be a definite relationship between these abilities for grade six students. The descrepant results for the grade seven students may be explained by this specific sample which did not perform as anticipated for the reasons pointed out earlier.

Null Hypothesis V

There is no significant relationship between the scores for students in grade six or students in grade seven on the Sentence Interpretation Test and

a) I.Q.

b) chronological age

The null hypothesis must be rejected for grade six subjects for I.Q., since Pearson product-moment correlations reached the level of significance set for the study. It cannot be rejected for chronological age for that group or for either I.Q. or chronological age for the grade seven group.

It would appear that when students make serious attempts to identify ambiguity in sentences (as with Little's grade five sample and the grade six subjects), intellectual ability may contribute to some degree to the relative success of that attempt. The lack of significant correlations between chronological age and the language tests of this study was not unexpected because of the small range of ages within either of the grade groups. A much larger sample covering

wider age range would be necessary to adequately test the relationships of chronological age and scores on the SIT and CAT.

Null Hypothesis VI

There is no significant relationship between the scores for students in grade six or students in grade seven on the Contextual Ambiguity Test and

a) I.Q.

b) chronological age

The null hypothesis cannot be rejected for the variable chronological age for either of the grades, since all correlations were close to zero. It also cannot be rejected for the variable I.Q. for grade seven students. However, for grade six students the null hypothesis is rejected for the relationship between I.Q. and the total CAT. The relationship between chronological age and scores on the CAT is explained again by the small range of ages within each grade, and the lack of relationship between grade seven scores and the variable I.Q. is explained by the atypical nature of this particular grade seven sample.

Null Hypothesis VII

There is no significant difference between boys and girls in grade six and grade seven in their ability to

a) identify ambiguity

b) disambiguate ambiguous sentences in context

Girls in both grades attained higher group scores than boys on all subtests except one of the SIT and CAT. However, since none of the differences reached the level of significance set for rejection of the null hypotheses of this study, the hypothesis cannot be rejected. The findings followed the same trend as Little's (1972) and earlier studies (Carroll, 1966; Balow, 1963; Weintraub, 1966; and Fagan, 1969), but as none of the observed differences was significant, they are not as conclusive.

Null Hypothesis VIII

There is no significant difference for students in grade six or grade seven among scores on unambiguous sentences, sentences with surface structure ambiguity, and sentences with underlying structure ambiguity of the Sentence Interpretation Test.

Tests on the differences between scores for the various combinations of three kinds of sentences on the SIT revealed that students in both grades performed in a significantly different way on each of the sets of sentence structures, and consequently the null hypothesis must be rejected.

The order of difficulty for both grades was the same as for Little's (1972) grade five subjects on the test: unambiguous sentences were easier to identify than surface structure ambiguities, and surface structure ambiguities were easier to identify than underlying structure ambiguities.

The order of difficulty between the last two sentence structure categories was the same as for adults as indicated by their perception time in identifying the two meanings of structurally ambiguous sentences (MacKay and Bever, 1967) and for adults who completed sentence fragments by orally producing complete sentences (MacKay, 1966). Although these two studies were quite unlike the present one in design and subjects studied, there is a similarity in the tasks which indicate that subjects perform better when they have to cope with surface structure ambiguities than with underlying structure ambiguities.

Null Hypothesis IX

There is no significant difference for students in grade six and students in grade seven between scores on passages containing surface structure ambiguity and passages containing underlying structure ambiguity on the Contextual Ambiguity Test.

This null hypothesis must be rejected for both grades on the basis of analyses of variance between the scores for the two types of ambiguous structures found in the test. The order of difficulty for both grades was reversed from what it had been found to be for the SIT. When constrained by a contextual passage, underlying structure ambiguities are more easily disambiguated than surface structure ambiguities. This order of difficulty was the same as that which occurred in Jurgens' (1971) study. Jurgens was measuring the time

which students took to select interpretative sentences as possible paraphrases of ambiguous sentences. Interpreting her findings, she suggested that because the identification of the meanings required the individual to regroup adjacent words, he essentially had to reprocess the sentence in order to test out alternative sentences as interpretations. Something similar may occur when the individual must test out interpretative sentences for the meaning allowed by the contextual paragraph, since the surface structure cues from the paragraph can only operate on the grouping of the words rather than on a logical relationship which exists between them.

Various researchers have put forth different suggestions as to how ambiguous sentences are processed. Possibly only one meaning of an ambiguous sentence is processed until it is found to be inadequate (Foss, Bever, and Silver, 1968), or both meanings are processed but one is suppressed by context (MacKay, 1970). The results of this study are unable to clarify the processing issue, but it appears that different models may be required to represent the processing of the two different types of structural ambiguities when they appear in context.

Summary

Because of the performance of the grade seven students on the Sentence Interpretation Test and the Contextual Ambiguity Test, the data could not be analyzed in terms of the

original purpose and design of this study. Nonetheless, analysis of the data carried out for both grades provided evidence for the following generalizations.

The task of dealing with ambiguous sentence structures in isolated sentences (identifying ambiguity as defined in this study) and in context (i.e. disambiguating ambiguous sentence strcutures) appear to be different and may reflect a difference in the processing required for each task.

Underlying structure ambiguities and surface structure ambiguities are different in nature as students operate on them regardless of whether they occur in isolated sentences or in the environment of a constraining paragraph.

There is evidence of developmental acquisition of the skill or skills involved in identifying ambiguity between grades five and six, but the development appears to be occurring at a fairly slow pace. Additional research will be required with students in grade seven and above to confirm the presence and rate of the developmental acquisition of the skills identified in the Sentence Interpretation Test and the Contextual Ambiguity Test.

A relationship appears to exist between the skills tested on the SIT and CAT and reading comprehension ability, but a clear pattern has not emerged regarding the relationships between inferential and literal reading comprehension abilities and the various abilities required by surface and underlying structure ambiguities within the two test situations.

A relationship also appears to exist [redacted] I.Q., reading comprehension ability, and the language skills tested.

III. IMPLICATIONS FOR EDUCATION

The information generated by the present study has the following implications for education:

1. There is some evidence that the presence of ambiguous structures within paragraphs, as can be found in the reading material of elementary and junior high students, may have a detrimental effect on reading comprehension for students in grade six and, possibly, grade seven. If this is the case, then teachers of all subjects which involve reading materials as part of the instructional framework need to be alert to the possible misinterpretations which students may be subject to and to ensure through discussion that the correct interpretation has been made by all students. Students who are not high achievers in reading and boys may particularly need this additional attention from their subject area teachers. Care should also be taken to see that teacher-prepared examination materials test the facts and concepts of the content material rather than the ability to identify or disambiguate structural ambiguity.

2. In order for teachers of subject areas which require reading to have the awareness required to implement the ideas from the first implication, they should have as part of their professional training preparation in under-

standing the reading processes, methods for utilizing reading materials for instructional purposes and methods for increasing reading skills, including reading comprehension. At present many secondary education (junior and senior high) teachers have no preparation in the reading area, although they utilize reading materials such as textbooks, library facilities, and examinations to a great extent.

It is equally important that the preparation in reading for teachers include sufficient linguistic background so that they become aware that language includes structures as conveying and sometimes interfering with the communication of information. These structures are flexible and can be manipulated, but students must be shown how to cope with them to get the most meaning from what they read.

3. In view of the relationship which appears to exist between the language competence represented by the abilities to identify and disambiguate ambiguous structures and reading comprehension, then direct instruction in the skills involved in the SIT and CAT could have the effect of raising general language competence and increasing reading comprehension ability. This instruction should focus not only on the potential difficulties which may be inherent in ambiguous structures, but also on the flexibility of conveying the same meaning through paraphrase, whether or not ambiguity is present and regardless of the changes in word order or appearances of words by means of prefixes and suffixes.

4. Refinement of tests similar to the SIT and CAT could lead to diagnostic instruments which would indicate a particular language competency of individual students. If the developmental aspects of the skills measured by these tasks are pursued, then the teacher or clinician could use the information to determine areas of weakness for individual students when compared to their peers and to identify specific skills for which instruction is needed.

5. Grade seven students appear to have lost a certain amount of enthusiasm for learning and are motivated only by marks and peer pressure. While this may not surprise the teachers of junior high students, it has implications for researchers who wish to use subjects from the junior high grades. Other motivators such as competition or material rewards may have to be used.

IV. SUGGESTIONS FOR FURTHER RESEARCH

The findings and limitations of the present study suggest the need for further research in the following areas:

1. Perhaps the most obvious suggestion is that the study be conducted with another grade seven sample to determine the relationship between the language skills measured by the tests of the experiment over various grade levels. Care must be taken to ensure that the grade seven sample selected are performing in keeping with their abilities.

2. The question of the relationship between the language skills measured by the SIT and CAT and reading comprehension needs to be pursued further, especially to determine whether inferential or literal reading comprehension are more directly related to any of the skills measured by the two language tests.

3. Research should be carried out with children from grades five on to determine what effects, if any, direct instruction in the abilities to identify and disambiguate structural ambiguities may have on reading comprehension abilities.

4. Researchers must continue to pursue testable models of the processing of ambiguity which may account for the differences which appear to exist between underlying structure and surface structure ambiguities.

V. CONCLUDING STATEMENT

This study has found a continuing development of the ability to identify structural ambiguity in sentences between students in grade five and grade six. This ability, and the ability to disambiguate structurally ambiguous sentences when constrained by a paragraph, have not fully developed in grade six students and appear to relate significantly to reading comprehension ability. High readers and girls in grade six were better able to identify and disambiguate ambiguity than were low readers and boys.

The performance of the grade seven group included in the sample did not appear to represent their actual abilities for the language skills measured by the language tests. Thus it is unwise to draw conclusions from the grade seven data of this study.

BIBLIOGRAPHY

Balow, I.H. Sex differences in first grade reading. Elementary English, 1963, 40, 303-306, 320.

Bormuth, J.R. et al. Children's comprehension of between- and within- sentence syntactic structures. Journal of Educational Psychology, 1970, 61, (5), 349-357.

Bormuth, J.R. Readability: a new approach. Reading Research Quarterly, 1966, 1, (3), 79-132.

Carey, P.W., J. Mehler, and T.G. Bever. Judging the veracity of ambiguous sentences. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 243-254.

Carroll, J.B. Language development. In C.W. Harris (ed) Encyclopedia of Educational Research. New York: The MacMillan Company, 1966, 744-752.

Carroll, J.B. et al. Word frequency book. New York: American Heritage Publishing Co. Inc., 1971

Chomsky, N. Syntactic structures. The Hague: Mouton & Co., 1957.

Coleman, E.B. The comprehensibility of several grammatical transformations. Journal of Applied Psychology, 1964, 48, 186-190.

Coleman, E.B. Improving comprehensibility by shortening sentences. Journal of Applied Psychology, 1962, 44, 131-134.

Cossitt, T.G. A linguistic analysis of social studies texts. Unpublished master's thesis, University of Alberta, 1966.

Dale, E. and J. Chall. A formula for predicting readability. Educational Research Bulletin, 1948, 27, 11-20.

Denner, B. Representational and syntactic competences of problem readers. Child Development, 1970, 41, (3), 881-887.

Fagan, W.T. An investigation into the relationship between reading difficulties and the number and types of sentence transformations. Unpublished doctoral dissertation, University of Alberta, 1969.

Forster, K.I. and L.A. Ryder. Perceiving the structure and meaning of sentences. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 107-115, 285-296.

Foss, D.J. Some effects of ambiguity upon sentence comprehension. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 699-706.

Foss, D.J., T.G. Bever and M. Silver. The comprehension and verification of ambiguous sentences. Perception and Psychophysics, 1968, 4, 304-306.

Garrett, M.L. Does ambiguity complicate the perception of sentences? In G.B. Flores D'Arcais and W.J.M. Levelt (eds.) Advances in Psycholinguistics. London: North-Holland Publishing Company, 1970, 48-60.

Gibbons, H.D. Reading and sentence elements. Elementary English Review, 1941, 18, 42-46.

Gray, W.S. and B.E. Leary. What makes a book readable. Chicago: University of Chicago Press, 1935.

Halliday, M.A.K. Speech and situation. In Some aspects of oracy. National Association of Teachers of English Bulletin, 1965, 2, 14-17.

Harris, Z. Papers in structural and transformational linguistics. Dordrecht-Holland: D. Reidel Publishing Co., 1970.

Helmstadter, G.C. Research concepts in human behavior. New York: Appleton-Century-Crofts, 1970.

Jacobs, R.A. and P.S. Rosenbaum. English transformational grammar. Waltham, Massachussetts: Blaisdell Publishing Company, 1968.

Jenkinson, M.D. Basic elements of reading comprehension. Paper presented at the Second World Congress on Reading, Copenhagen, 1968.

Jurgens, J.M. Perception of lexical and structural ambiguity by junior and senior high school students. Unpublished doctoral dissertation, George Peabody College, 1971.

Kessel, F.S. The role of syntax in children's comprehension from ages six to twelve. Monograph of the society for research in child development, 1970, 35, (6), Serial No. 139.

[redacted] P.S. An investigation into relationship between structural ambiguity and reading comprehension. Unpublished [redacted]ter's thesis, Univeristy of Alberta, 1972.

Lorge, I. Predicting readability. Teachers College Record, 1944, 45, 404-419.

MacKay, D.G. To end ambiguous sentences. Perception and Psychophysics, 1966, 1, 426-436.

MacKay, D.G. Mental diplopia: Toward a model of speech perception at the semantic level. In G.B. Flores D'Arcais and W.J.M. Levelt (eds.) Advances in Psycholinguistics. London: North-Holland Publishing Company, 1970.

MacKay, D.G. and T.G. Bever. In search of ambiguity. Perception and Psychophysics, 1967, 2, 193-200.

Menyuk, P. The acquisition and development of language. Englewood Cliffs, N.J.: Prentice Hall, 1971.

O'Donnell, R. Awareness of grammatical structure and reading comprehension. High School Journal, 1962, 45, 184-188.

Prideaux, G.D. Surface structure resolution of structural ambiguity. Paper presented to the Annual Meeting of the Alberta Linguistic Association, Banff, Alberta, 1972.

Robertson, J.E. An investigation of pupil understanding of connectives in reading. Unpublished doctoral dissertation, University of Alberta, 1966.

Rommetveit, R. Words, meanings and messages. New York: Academic Press, 1968.

Ruddell, R.B. An investigation of the effect of the similarity of oral and written patterns of language structure on reading comprehension. Unpublished doctoral dissertation, Indiana University, 1963.

Simons, H.D. Linguistic skills and reading comprehension. U.S. Department of Health, Education, and Welfare, Office of Education Bureau of Research, Final Report, Project No. 9-A-078, 1970.

Spache, G.D. A new readability formula for primary grade reading materials. Elementary School Journal, 1953, 53, 410-413.

Strom, J.M. Does knowledge of grammar improve reading? English Journal, 1956, 44, 129-133.

Tatham, S.M. Reading comprehension of materials written with select oral language patterns: a study at grades two and four. Reading Research Quarterly, 1970, 5, (3), 427-451.

Taylor, W.L. Cloze procedure: a new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-433.

Vogel, M. and C.W. Washburne. An objective method of determining grade placement of children's reading material. Elementary School Journal, 1928, 26, 373-381.

Weintraub, S. Sex differences in reading achievement. The Reading Teacher, 1966, 20, 155-165.

APPENDIX A

LANGUAGE QUESTIONNAIRE

APPENDIX A

LANGUAGE QUESTIONNAIRE

NAME: _____ CLASS_____

TEACHER _____

Did you speak English when you started school?

Yes _____ No _____

Did you speak any other language when you started school?

Yes _____ No _____

If so, which language? _____

APPENDIX B

RATIONALE FOR SENTENCE INTERPRETATION TEST

RATIONALE FOR SENTENCE INTERPRETATION

II. TESTING INSTRUMENTS

1. Sentence Interpretation Test

The Sentence Interpretation Test, (SIT), used in this study to measure the ability of children to identify the meanings of structurally ambiguous or unambiguous sentences of written English, was constructed by the investigator. It consisted of forty lead sentences: ten with

surface structure ambiguity, ten with underlying structure ambiguity, and twenty which were unambiguous. For each of these lead sentences three interpretative sentences were constructed, one, two, or all three of which gave a meaning of the lead sentence. An example of a complete test item is given in Figure 1.

	GIVES A MEANING	DOES NOT GIVE A MEANING
BOYS LIKE ICE CREAM BETTER THAN GIRLS.		
(a) It is ice cream that boys like better than they like girls.	_____	_____
(b) Boys like ice cream better than girls like boys.	_____	_____
(c) Boys like ice cream better than girls like ice cream.	_____	_____

Fig. 1. Sample SIT Item

Sentences with lexical ambiguity were not included in the test because of the findings by Kessel (1970), later confirmed by Jurgens (1971), that the perception of lexical ambiguity is highly dependent on the actual lexical item containing the ambiguity and thus depends largely on the vocabulary knowledge of the individual, and for this reason cannot be considered on the same developmental continuum as the ability to perceive structural ambiguity.

Construction of the Lead Sentences for the SIT

The construction of the lead sentences for the SIT was based upon an analysis of the syntactic structures that occurred in the

structurally ambiguous sentences used by MacKay (1966), MacKay and Bever (1967), and Jurgens (1971). This analysis revealed that the types of syntactic structures in which the ambiguity was located differed absolutely between those sentences classified as containing surface structure ambiguity and those classified as containing underlying structure ambiguity. To the extent that these researchers did not make provision for these types of structure to occur equally in all types of sentences included in their tests, it was felt that their conclusions about the effect of structural complexity on the perception of ambiguity were invalid.

Five main structures for both types of structurally ambiguous sentences were identified. These were as follows:

(a) Surface Structure Ambiguities

1. Adjective + Noun + Noun - where the element Noun + Noun may be interpreted as a compound noun, in which case the Adjective modifies the second Noun in the compound noun, or where both Nouns are distinct, in which case the Adjective modifies the first Noun. e.g., He was an American art expert.

2. Adverb/Adjective - where one word which may function as either an Adverb or an Adjective may be interpreted alternately. e.g., The blue dress particularly interested her.

3. Prepositional Phrase - where the Prepositional Phrase may be interpreted as modifying either a preceding noun or a preceding verb. e.g., He painted the picture on the patio.

4. Adjective + Noun_1 + and + Noun_2 - where the Adjective may

be interpreted as modifying only $Noun_1$ or, by a common-elements deletion transformation, as modifying both $Noun_1$ and $Noun_2$. e.g., Little boys and girls enjoy watching fireworks.

5. $Noun_1$ + $Noun_2$ - where one Noun immediately following another in a terminal string may be interpreted as either a compound noun or two separate nouns. e.g., He told her baby stories.

(b) Underlying Structure Ambiguities

1. Infinitive - where the infinitive may be interpreted as "transitive" with an unspecified object, or as "intransitive" with "be" deleted. e.g., The lamb is too hot to eat.

2. Verb+ing + Noun - where Verb+ing may be interpreted as part of a verbal or as an adjective modifying the following Noun. e.g., He disliked visiting relatives.

3. Genitive Construction - where the genitive may be interpreted as deriving from an underlying structure of the form That + Determiner + Noun + Verb + Something or from an underlying structure of the form That + Determiner + Noun + be + Verb. e.g., The manager's selection was announced.

4. Infinitive + Verb+ing - where Verb+ing may be interpreted either as part of the verbal containing the Infinitive or as a nominalization which functions as the object of the Infinitive. e.g., The police were asked to stop drinking.

5. Comparative Deletion - where the deleted elements in a comparison may be interpreted as being either the Subject + Verb of the sentence or the Verb + Object of the sentence. e.g., Boys like tennis better than girls.

As a basis for constructing the lead sentences, those structures characteristic of surface structure ambiguity were randomly paired with those characteristic of underlying structure ambiguity. These pairings are shown in Figure 2.

Pairing	Structure Characteristic of Surface Structure Ambiguity	Structure Characteristic of Underlying Structure Ambiguity
1.	Adj + N + N	Infinitive
2.	Adv/Adj	Ving + N
3.	Prep Phrase	Genitive
4.	Adj + N_1 + and + N_2	Infinitive + V ing
5.	N_1 + N_2	Comparative Deletion

Fig. 2. Pairings of Structures for Construction of the SIT

Forty lead sentences were then constructed such that there were eight lots of five sentences each. Two of the eight lots were surface structure ambiguities, two were underlying structure ambiguities, and four were unambiguous. These were designated as Types 1 to 8 according to the nature of the structures that they contained. The construction of these types is described in detail below using Pairing 1. (Fig. 2) as the basis for example.

The five Type 1 sentences (surface structure ambiguities) were constructed so that each sentence in this Type contained one of the structure pairings shown in Fig. 2. In this Type the structure characteristic of surface structure ambiguity was used ambiguously

and the structure characteristic of underlying structure ambi used unambiguously. For example, the sentence He went to fetch the red crayon box in which the structure Adj + N + N is ambiguous but the infinitive structure is unambiguous.

Type 2 sentences (underlying structure ambiguities) were likewise constructed so that each sentence contained one of the structure pairings. However, this Type differed in that the ambiguity resided in that structure characteristic of underlying structure ambiguity and the structure characteristic of surface structure ambiguity was used unambiguously. For example, the sentence The young science teacher is the one to ask, in which the the infinitive structure is ambiguous but the Adj + N + N structure is used unambiguously.

Type 3 and Type 4 sentences (unambiguous) were constructed to balance Type 1 and Type 2 sentences respectively in terms of syntactic complexity. That is, each one of the sentences in Type 3 was constructed to contain exactly the same structures as each one of the sentences in Type 1, and likewise each one of the sentences in Type 4 was constructed to contain exactly the same structures as each one of the sentences in Type 2. These Types differed from Types 1 and 2 only in that all structures were used unambiguously.

For example, Type 3 sentence He wanted to find the front door key (compare Type 1 He went to fetch the red crayon box), and Type 4 sentence The white race horse was the first to finish (compare Type 2 The young science teacher is the one to ask).

Thus, Types 1, 2, 3 and 4 all contained the same basic syntactic

structures.

Type 5 sentences (surface structure ambiguities) were constructed so that each sentence in this type contained an ambiguous usage of one of the five structures characteristic of surface structure ambiguity. However, these sentences did not contain the paired structure shown in Figure 2, and no control was placed upon the structure of the remainder of each sentence except that it be unambiguous. For example, the sentence A bitter bear trap was sold to the hunter contains the Adj + N + N structure which is used ambiguously (compare Type 1 He went to fetch the red crayon box) but as has been indicated, the sentences of Type 5 did not contain the pairing structure of Type 1. In this case, therefore, the infinitive structure of Type 1 was not repeated in Type 5.

Type 6 sentences (underlying structure ambiguities) were constructed in the same way as Type 5 sentences except that they contained an ambiguous usage of one of the five structures characteristic of underlying structure ambiguity rather than of surface structure ambiguity. For example, the sentence The hunter was too far away to see (compare Type 2 The young science teacher is the one to ask).

Type 7 sentences (unambiguous) were constructed to contain identical structures to those used in Type 5 sentences except that all structures were used unambiguously. The syntactic interpretation used for the structure characteristic of surface structure ambiguity was the alternate of that used when constructing Type 3 sentences. For example, the sentence A new wrist watch was given to the winner (compare Type 5

A black bear trap was sold to the hunter).

Type 8 sentences (unambiguous) were constructed to contain identical structures to those used in Type 6 sentences, again with the exception that all structures were used unambiguously. Likewise, the syntactic interpretation used for the structure characteristic of underlying structure ambiguity was the alternate of that used when constructing Type 4 sentences. For example, the sentence The box was too high up to reach (compare Type 6 The hunter was too far away to see).

A summary of the basic structural design of these eight types of sentences and the SIT item numbers corresponding to each type are contained in Figure 3.

Sentence Type	Sentence Structure Design	Test Item No.s
1	surface structure ambiguity + unambiguous structure characteristic of underlying structure ambiguity	10,26,27,33,37
2	underlying structure ambiguity + unambiguous structure characteristic of surface structure ambiguity	4,16,18,39,40
3	unambiguous instances of these structures occurring in Type 1 sentences	3,6,11,21,35
4	unambiguous instances of these structures occurring in Type 2 sentences	5,7,15,25,30
5	surface structure ambiguity + optional structure	9,13,17,19,20
6	underlying structure ambiguity + optional structure	1,2,22,28,31
7	unambiguous instances of those structures occurring in Type 5 sentences	8,12,23,32,38
8	unambiguous instances of those structures occurring in Type 6 sentences	14,24,29,34,36

Fig. 3. Sentence Types in the SIT

Other considerations that affected the construction of these lead sentences were sentence length, vocabulary, grammatical and semantic acceptability, and the consistency with which these sentences could be classified as ambiguous or unambiguous by mature native speakers of English.

a. Sentence Length - In view of the importance attached to sentence length in studies of the readability of written language, it was considered necessary to maintain a consistent sentence length for each lead sentence. Thus the precedent of MacKay and Bever (1967) was followed in limiting each sentence to eight words (plus or minus one).

b. Vocabulary - In order that the vocabulary used in the sentences constructed could be considered as within the reading vocabulary of grade five students, only words listed in Carroll's (1971) Word Frequency Book as occurring in the reading material of grade four students or below were used.

c. Grammatical and Semantic Acceptability and Consistency of Classification - a panel composed of twenty graduate students and faculty members was asked to classify each of the lead sentences as ambiguous or unambiguous and to comment on the grammatical and semantic acceptability of these sentences. Any ambiguous sentence that was not classified as such by at least fifteen of the panel and any unambiguous sentence that was classified as ambiguous by any one member of the panel were revised or replaced. Revised and replacement sentences were again submitted for judgement until the investigator was reasonably certain that the lead sentences adequately fitted the category assigned to them

and were grammatically and semantically acceptable.

Construction of the Interpretative Sentences for the SIT

Three other sentences were then constructed for each lead sentence such that either one, two, or all three of these sentences gave a paraphrased meaning of the lead sentence. For all ambiguous lead sentences, two of the three sentences gave a meaning, representing both interpretations of the ambiguity. To avoid any overt pattern to the number of correct responses for each item, this number was varied for the unambiguous sentences. Thus, for the unambiguous lead sentences constructed to parallel the syntactic complexity of those lead sentences containing surface structure ambiguity, four were randomly assigned to have only one of the three interpretative sentences give a meaning, four more were assigned to have all three of the interpretative sentences give a meaning, and the remaining two were assigned to have two of the three interpretative sentences give a meaning. The same procedure was followed for the unambiguous sentences constructed to parallel the syntactic complexity of those lead sentences containing underlying structure ambiguity. Each of these interpretative sentences was constructed such that the least possible change was made in the wording of the lead sentence to represent the required meaning. In no instance were any content words introduced into the interpretative sentences that did not occur in the lead sentence.

The ordering of the interpretative sentences was randomized for each lead sentence and the lead sentences themselves were ran-

domized with the exception that no two pairings of the same type, as represented in Figure 2, were permitted to immediately follow one another. This precaution was taken as it was felt that two sentences of similar structure occurring together might influence the interpretation of each other.

Instructions for the SIT

The instructions for the test contained one example of an item with surface structure ambiguity, one of an item with underlying structure ambiguity and one that was unambiguous. The students were instructed to read each lead sentence carefully, read the interpretative sentences and indicate which of the latter gave a meaning of the lead sentence by placing a check (✓) by each interpretative sentence under a column headed "GIVES A MEANING" or under a column headed "DOES NOT GIVE A MEANING".

The complete test with its instructions is contained in Appendix A.

Validity of the SIT

Helmstadter (1970) states that "in the original writing of items, face validity is about all there is to rely upon (p. 298)."

Face validity for the SIT as a measure of the ability to identify the meanings of structurally ambiguous or unambiguous sentences of written English is claimed on the basis of the following considerations:

1. That structurally ambiguous sentences differ from unambiguous

sentences. This was determined by a structural analysis of the lead sentences in terms of transformational generative grammar and was confirmed by submitting all sentences to a panel of mature native speakers of English for judgement. Also, the results of the pilot study support the evidence of MacKay (1966) and Jurgens (1971) that structurally ambiguous sentences are more difficult to interpret than are unambiguous sentences.

2. That sentences with surface structure ambiguity differ from sentences with underlying structure ambiguity. Although Prideaux (1972) has shown that both types of structural ambiguity are resolvable at the level of surface structure, the research of MacKay (1966), MacKay and Bever (1967), Kessel (1970) and Jurgens (1971) has shown that an individual's perception of these types of sentences differs. Results of the pilot study also suggested that children's ability to identify the meanings of surface structure ambiguities differs from their ability to identify the meanings of underlying structure ambiguities.

3. That sentences that are paraphrases of each other have the same deep structure. This is a basic fact of language as described by the competence theory of transformational generative grammar and unless this theory is proven wrong, it seems reasonable to accept the use of paraphrases as interpretations of the lead sentences in the test.

4. That the students choice of the correct interpretative sentences for the lead sentences indicates that they have recovered the

deep structures and thus the meanings of that lead sentence. Simons (1970) has provided ample evidence that the recovery of deep structure is necessary to determine whether or not sentences are paraphrases of one another. If this is the case, and the theory of transformational generative grammar would again support this contention, then the use of paraphrases of the lead sentences to measure the students' ability to identify the meanings of the lead sentences would seem to be valid.

5. That vocabulary was carefully controlled, and that readers of average ability were selected for the sample, would indicate that the lexical items used in the test were well within the range of comprehensibility of average grade five students, and thus the test was not measuring the variables of word recognition and word identification that are associated with reading comprehension. This was borne out by the pilot study and interviews conducted with certain students in the sample after they had taken the test.

6. That the sentences were grammatically and semantically acceptable was adjudged by a panel of mature, educated, native speakers of English and again borne out by the pilot study and interviews.

Reliability of the SIT

On the basis of data collected on sixty grade five students, the reliability of the SIT was calculated by using the technique of split-half reliability. The items were so divided that each half of the test contained equal numbers of sentences that were unambiguous, that contained surface structure ambiguity, and that contained underlying structure ambiguity. The resultant correlation of .722 was

corrected by the Spearman-Brown prophecy formula to determine the reliability of the entire test. The reliability of the SIT is .839 (pp. 40-50).

APPENDIX C

SENTENCE INTERPRETATION TEST (SIT)

NAME: _____ CLASS _____

SENTENCE INTERPRETATION TEST

INSTRUCTIONS: Normally, the sentences that you read have only one meaning, but sometimes they can have more than one meaning. This test is designed to see how well you can understand the meaning or meanings of a sentence. In each question a lead sentence is typed in capital letters and underlined. Below this sentence are three more sentences labelled a), b), and c). One or more of these sentences may give a meaning which is the same as the meaning of the underlined sentence. Take each one in turn, compare it to the underlined sentence, and decide whether it gives a meaning which is the same as the meaning of the underlined sentence or not. For each one that does, put a check under the column GIVES A MEANING, and for each one that doesn't, put a check under the column marked DOES NOT GIVE A MEANING. Look at these examples:

		GIVES A MEANING	DOES NOT GIVE A MEANING
Ex. 1:	HE KNEW THAT BURNING PAPER COULD BE DANGEROUS.		
	a) He knew that burning could be dangerous to paper.	____	✓
	b) He knew that paper which was burning could be dangerous.	✓	____
	c) He knew that it could be dangerous to burn paper.	✓	____

	GIVES A MEANING	DOES NOT GIVE A MEANING
Ex. 2: THE ANGRY CAT WAS CHASING THE DOG.		
a) The angry dog was being chased by the cat.	____	✓
b) The dog was chasing the angry cat.	____	✓
c) It was the angry cat that was chasing the dog.	✓	____
Ex. 3: SAM PAINTED THE PICTURE IN THE KITCHEN.		
a) The picture that Sam painted was in the kitchen.	✓	____
b) What Sam painted was the picture of the kitchen.	____	✓
c) It was in the kitchen that Sam painted the picture.	✓	____

The lead sentence in Ex. 3 could have two meanings - either that the picture was in the kitchen, or that Sam did the painting in the kitchen. Thus you should have put a check by a) and c) under GIVES A MEANING, and a check by b) under DOES NOT GIVE A MEANING.

Remember that in the test sometimes only one, sometimes two, and sometimes all three sentences will give a meaning for the underlined sentence. Use only the information contained in the sentences to make your decisions.

	GIVES A MEANING	DOES NOT GIVE A MEANING
1. THE HUNTER WAS TOO FAR AWAY TO SEE.		
a) The hunter was too far away to be seen.	____	____
b) The hunter was too far away from something to see it.	____	____
c) The hunter who was seen was too far away.	____	____

	GIVES A MEANING	DOES NOT GIVE A MEANING
2. THE CHOICE OF THE STUDENTS SURPRISED THE PARENTS.		
a) What surprised the parents was the students who were chosen.	_____	_____
b) It was the choice made by the parents that surprised the students.	_____	_____
c) The choice made by the students surprised the parents.	_____	_____
3. HE WANTED TO FIND THE FRONT DOOR KEY.		
a) It was the front door key that he wanted to find.	_____	_____
b) He wanted to find the front door for the key.	_____	_____
c) He wanted to find the key for the front door.	_____	_____
4. ONLY THOSE LADIES WHO LIKED GROWING FLOWERS CAME.		
a) Only those ladies who were like growing flowers came.	_____	_____
b) Only those ladies came who liked flowers that are growing.	_____	_____
c) It was only those ladies who liked to grow flowers that came.	_____	_____
5. THE TIGER'S ROAR WAS HEARD DURING THE NIGHT.		
a) It was heard that the tiger had roared during the night.	_____	_____
b) During the roar of the night the tiger was heard.	_____	_____
c) The roar of the tiger was heard during the night.	_____	_____

	GIVES A MEANING	DOES NOT GIVE A MEANING
6. YOUNG BILLY AND DAD LIKE TO GO FISHING.		
a) What young Billy likes is for Dad to go fishing.	_____	_____
b) Dad likes young Billy to go fishing.	_____	_____
c) It is young Billy and Dad who like to go fishing.	_____	_____
7. ONLY THOSE BOYS WHO WANTED WRITING PAPER STAYED.		
a) Only those boys who wanted to write a paper stayed.	_____	_____
b) It was only those boys who wanted paper for writing that stayed.	_____	_____
c) Only those boys stayed who wanted paper that was for writing.	_____	_____
8. SOME STORES SELL NEW TOYS AND NEW BOOKS.		
a) New toys and new books are sold by some stores.	_____	_____
b) Some stores sell new books and new toys.	_____	_____
c) It is new toys and new books that some stores sell.	_____	_____
9. A SMALL BOAT ENGINE WAS SOLD TO THE CAPTAIN.		
a) To the captain was sold an engine for a small boat.	_____	_____
b) A small engine for a boat was sold to the captain.	_____	_____
c) A small boat and an engine were sold to the captain.	_____	_____

	GIVES A MEANING	DOES NOT GIVE A MEANING
10. HE THREW HER DOG BISCUITS RATHER THAN CANDIES.		
a) It was dog biscuits that he threw her rather than candies.	____	____
b) He threw candies to her rather than dog biscuits.	____	____
c) He threw biscuits rather than candies to her dog.	____	____
11. IT FELL TO THE GROUND WITH A CRASH.		
a) It was with a crash that it fell to the ground.	____	____
b) It crashed and then it fell to the ground.	____	____
c) It fell with a crash to the ground.	____	____
12. HE TOLD HER BABY STORIES IN THE EVENING.		
a) Baby stories were what he told her in the evening.	____	____
b) She told him baby stories in the evening.	____	____
c) In the evening he told stories to her baby.	____	____
13. MOM LIKES SEA FOOD BETTER THAN MEAT.		
a) Meat is what Mom likes better than sea food.	____	____
b) It is meat that Mom likes better than she likes sea food.	____	____
c) Mom likes sea food better than she likes meat.	____	____

	GIVES A MEANING	DOES NOT GIVE A MEANING
14. THE YOUNG SCIENCE STUDENT IS THE ONE TO ASK.		
a) The young science student is the one who should do the asking.	____	____
b) The young science student was the one who asked.	____	____
c) The one who should be asked is the young science student.	____	____
15. SHE SPOKE TO THE BOY WITH A SMILE.		
a) She smiled and then spoke to the boy.	____	____
b) The boy with a smile was the one to whom she spoke.	____	____
c) She was smiling when she spoke to the boy.	____	____
16. THE NEW TEACHER AND TOM WANTED TO START SKATING.		
a) It was skating that the new teacher and Tom wanted to start.	____	____
b) The new teacher and Tom were going to skate and they wanted to start.	____	____
c) The new teacher wanted Tom to start skating.	____	____
17. MOST PEOPLE WEAR CLEAN SOCKS AND SHOES.		
a) Most people wear shoes and clean socks.	____	____
b) Most people wear socks and clean shoes.	____	____
c) It is clean socks and clean shoes that most people wear.	____	____

	GIVES A MEANING	DOES NOT GIVE A MEANING
18. HER BEST FRIEND HAD LEFT EARLY TONIGHT.		
a) Tonight her friend had best leave early.	_____	_____
b) Early tonight was when her best friend had left.	_____	_____
c) Tonight her best friend had left early.	_____	_____
19. SHE WANTED A CANDY BAR MORE THAN A POP.		
a) She wanted a bar of candy more than she wanted a pop.	_____	_____
b) It was a candy bar that she wanted more than a pop.	_____	_____
c) What she wanted more than a pop was a candy bar.	_____	_____
20. THE DOG FOUND TOM MORE QUICKLY THAN BOBBY.		
a) It was Bobby who was found by the dog more quickly than Tom.	_____	_____
b) The dog found Tom more quickly than the dog found Bobby.	_____	_____
c) The dog found Tom more quickly than Bobby found Tom.	_____	_____
21. A NEW WRIST WATCH WAS GIVEN TO THE WINNER.		
a) The winner was given a new wrist watch.	_____	_____
b) A wrist watch that was new was given to the winner.	_____	_____
c) What was given to the winner was a new wrist watch.	_____	_____

	GIVES A MEANING	DOES NOT GIVE A MEANING
22. LITTLE JACK AND GRANDFATHER DECIDED TO GO CAMPING.		
a) It was decided by little Jack and Grandfather that they go camping.	____	____
b) To go camping was what little Jack and Grandfather decided.	____	____
c) Grandfather and little Jack decided to go camping.	____	____
23. HE WENT TO FETCH THE RED PAINT BOX.		
a) He went to fetch the box of red paint.	____	____
b) It was the red box of paint that he went to fetch.	____	____
c) He went to fetch the red paint for the box.	____	____
24. THE DANCING TEENAGERS KNEW HOW GOOD MUSIC SOUNDED.		
a) The dancing teenagers knew how to sound good at music.	____	____
b) The dancing teenagers knew how good the sound of music was.	____	____
c) The teenagers who were dancing knew how music which is good sounded.	____	____
25. THE PRINCIPAL ASKED THE TEACHERS TO STOP SMOKING.		
a) The principal asked the teachers to stop others from smoking.	____	____
b) Other people were smoking and the teachers asked the principal to stop them.	____	____
c) The teachers were smoking and the principal asked them to stop.	____	____

	GIVES A MEANING	DOES NOT GIVE A MEANING
26. THE CAT DRINKS MILK FASTER THAN THE BABY.		
a) It is the cat that drinks milk faster than the baby.	_____	_____
b) The cat drinks milk faster than the baby drinks milk.	_____	_____
c) Milk is drunk faster by the cat than by the baby.	_____	_____
27. THE WHITE RACE HORSE WAS THE FIRST TO FINISH.		
a) The white race horse was the one that finished first.	_____	_____
b) The first to finish was the race horse that was white.	_____	_____
c) It was the white race horse that was first to finish.	_____	_____
28. MOST PEOPLE WHO SHOP LIKE PLEASING SALESMEN.		
a) Salesman who are pleasing are liked by most people who shop.	_____	_____
b) Most people who shop like to please salesmen.	_____	_____
c) Salesmen like pleasing most people who shop.	_____	_____
29. MY OLD UNCLE HAD WORKED HARD TODAY.		
a) It was my old uncle who had worked hard today.	_____	_____
b) Today was when my old uncle had worked hard.	_____	_____
c) Today my old uncle had worked hard.	_____	_____
30. LITTLE CATS AND DOGS LIKE TO GO EXPLORING.		
a) Cats and little dogs like to go exploring.	_____	_____
b) Little cats and little dogs like to go exploring.	_____	_____
c) It is dogs and little cats that like to go exploring.	_____	_____

	GIVES A MEANING	DOES NOT GIVE A MEANING
31. THE DISCOVERY OF THE ISLAND PLEASED THE KING.		
a) That the island was discovered pleased the king.	_____	_____
b) The king was pleased by the discovery of the island.	_____	_____
c) What pleased the king was the discovery of the island.	_____	_____
32. JOHN'S BALL BROKE THE WINDOW IN THE DOOR.		
a) The window in the door was broken by John's ball.	_____	_____
b) John's ball was broken by the window in the door.	_____	_____
c) The door by the window was broken by John's ball.	_____	_____
33. DAD TOOK THE BROTHERS TO WATCH BOXING.		
a) The brothers took Dad to watch boxing.	_____	_____
b) Boxing was what Dad took the brothers to watch.	_____	_____
c) The brothers were taken by Dad to watch boxing.	_____	_____
34. MICHAEL'S BROTHER WATCHED THE GAME ON THE PATIO.		
a) The game on the patio was watched by Michael's brother.	_____	_____
b) While on the patio Michael watched his brother's game.	_____	_____
c) While on the patio Michael's brother watched the game.	_____	_____

	GIVES A MEANING	DOES NOT GIVE A MEANING
35. MOM GAVE THE BOY MONEY FROM HER PURSE.		
a) The boy gave Mom money from her purse.	____	____
b) It was the boy who gave Mom money from her purse.	____	____
c) Mom gave money to the boy from her purse.	____	____
36. BOYS LIKE ICE CREAM BETTER THAN GIRLS.		
a) It is ice cream that boys like better than they like girls.	____	____
b) Boys like ice cream better than girls like boys.	____	____
c) Boys like ice cream better than girls like ice cream.	____	____
37. THE TEAM'S SELECTION WAS ANNOUNCED IN THE MORNING		
a) The morning's selection was announced by the team.	____	____
b) That the team was selected was announced in the morning.	____	____
c) The selection made by the team was announced in the morning.	____	____
38. THE TRAVELING SALESMAN ASKED HOW THE YOUNG LADY WAS.		
a) The traveling salesman was asked how the young lady was.	____	____
b) The traveling salesman asked how the lady who was young was.	____	____
c) The salesman who was traveling asked how young the lady was.	____	____

	GIVES A MEANING	DOES NOT GIVE A MEANING
39. THE LIMB WAS TOO HIGH UP TO GRAB.		
a) The limb was too high up to be grabbed.	_____	_____
b) The limb that was grabbed was too high up.	_____	_____
c) Someone was too high up to grab the limb.	_____	_____
40. MOST PEOPLE WHO TRAVEL LIKE EATING SNAILS.		
a) Snails like eating most people who travel.	_____	_____
b) Most people who travel like to eat snails.	_____	_____
c) It is most people who travel that eat like snails.	_____	_____

APPENDIX D

ITEM ANALYSIS OF (SIT)

ITEM ANALYSIS OF SIT

TEST Item No.	Biserial Correlation	Item Difficulty Index
1.	.479	.185
2.	.628	.251
3.	.724	.158
4.	.633	.268
5.	.080	.040
6.	.551	.099
7.	.478	.197
8.	.769	.317
9.	.366	.183
10.	.180	.090
11.	.387	.180
12.	.521	.257
13.	.286	.114
14.	1.120	.359
15.	1.194	.330
16.	.710	.345
17.	.602	.248
18.	.400	.160
19.	.399	.183
20.	.763	.331
21.	.652	.233
22.	.630	.266
23.	.579	.289
24.	.565	.280
25.	1.073	.296
26.	.679	.253
27.	.736	.285
28.	.660	.255
29.	.778	.321
30.	.708	.264
31.	.561	.280
32.	.214	.083
33.	.760	.228
34.	.291	.137
35.	.447	.111
36.	.659	.279
37.	.433	.183
38.	.276	.130
39.	.192	.034
40.	.261	.119

APPENDIX D

ITEM ANALYSIS OF (SIT)

APPENDIX E

CONTEXTUAL AMBIGUITY TEST (CAT)

NAME: _____ CLASS _____

CONTEXTUAL AMBIGUITY TEST

INSTRUCTIONS: This test is to find out if you understand the meaning of a sentence or group of words which occur in a paragraph. In the passages below, one of the sentences or groups of words is underlined. Below the passages are three sentences labelled a), b), and c). Some of these sentences give the same meaning as the meaning of the underlined words in the paragraph. Read the whole paragraph. Then read the sentence labelled a). Compare the meaning of this sentence with the meaning of the underlined words in that paragraph. If the sentence gives a meaning that is the same as the meaning of the underlined words in the paragraph, check GIVES A MEANING. If it does not give a meaning that is the same as the meaning of the underlined words in the paragraph, check DOES NOT GIVE A MEANING. Do the same for sentence b) and sentence c). Then go on to the next paragraph and do the same.

EXAMPLE: Sam was given a set of oil paints for his birthday. He immediately set up his easel in the livingroom and began to sketch the view from the front window. His mother found him there and began yelling about not getting paint on her new carpet. So <u>Sam painted the picture in the kitchen</u>, with frequent visits to the livingroom to check the details of the view.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Sam painted the picture that was in the kitchen.	____	✓
b) What Sam painted was the picture of the kitchen.	____	✓
c) It was in the kitchen that Sam painted the picture.	✓	____

1. George heard a shot from somewhere. He climbed a tree and scanned the horizon for the red of a hunter's shirt, but the hunter was too far away to see. George hoped that the yearling deer would stay in the bush where he had gone for safety. He prayed that the hunter would miss if he saw the deer.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) The hunter was too far away to be seen.	____	____
b) The hunter was too far away from something to see it.	____	____
c) The hunter who was seen was too far away.	____	____

2. The graduating class at Central School decided against having a dinner and dance with an expensive band. Instead they would have a casual picnic and entertainment by members of the class themselves. Everyone would donate the amount of money usually spent on a banquet and dance, and the money would be given to charity. When they learned these plans, the parents were surprised by the choice of the students.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) What surprised the parents was the students who were chosen.	_____	_____
b) It was the choice made by the parents that surprised the students.	_____	_____
c) The choice made by the students surprised the parents.	_____	_____

3. Last summer, the Ladies' Auxiliary gave a garden fair to raise money for Children's Hospital. They sold potted plants, hand-made gifts, and baked goods. However, the fair was a failure. Because of the name "garden fair", only those ladies who liked growing flowers came. They were disappointed that there were no speakers on how to prune rose bushes or when to fertilize petunia beds. They left without buying anything.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Only those ladies who were like growing flowers came.	_____	_____
b) Only those ladies came who liked flowers that are growing.	_____	_____
c) It was only those ladies who liked to grow flowers that came.	_____	_____

4. In the month of July, Captain Brown made three trips to Seaside Boat Equipment. In the first week, he bought some paint to paint his boathouse. The next week, he bought a medium-sized engine for his fishing boat. The engine was too powerful for the boat, however, and the following week a small boat engine was sold to the captain.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) To the captain was sold an engine for a small boat.	_____	_____
b) A small engine for a boat was sold to the captain.	_____	_____
c) A small boat and an engine were sold to the captain.	_____	_____

5. If you want a cat or a dog that will stay close to home and not go wandering about the streets, it is better to buy one that is as large as possible. Big cats and dogs seem to like to stay close to their home, but little cats and dogs like to go exploring.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Cats and little dogs like to go exploring.	_____	_____
b) Little cats and little dogs like to go exploring.	_____	_____
c) It is dogs and little cats that like to go exploring.	_____	_____

6. Mrs. Grant was always grateful to Uncle Dave for entertaining her baby, Paul, when he visited. Usually Paul went to the park with Uncle Dave in the afternoon. Later, when Mrs. Grant was fixing supper, he would play with the baby in the livingroom. Last time, he told her baby stories in the evening, while Mrs. Grant went out to do some shopping. He was a good babysitter.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Baby stories were what he told her in the evening.	_____	_____
b) She told him baby stories in the evening.	_____	_____
c) In the evening he told stories to her baby.	_____	_____

7. Have you ever wondered how two brown-eyed parents can produce a blue-eyed child? For the answer to that question, the young science student is the one to ask. Students in grade six science classes in our city have been doing controlled experiments with fruit flies and peas to discover how physical characteristics are passed from parents to offspring.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) The young science student is the one who should do the asking.	_____	_____
b) The one who should be asked is the young science student.	_____	_____
c) The young science student was the one who asked.	_____	_____

8. Yesterday, while Suzie was playing tennis, she hit the ball so hard that it went over the fence. It landed on the head of a boy who was playing marbles. The boy scrambled after the tennis ball, scowling angrily. He was about to fling it into a mud puddle when Suzie ran up to him. She spoke to the boy with a smile, and he handed her the ball.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) She was smiling when she spoke to the boy.	_____	_____
b) The boy with a smile was the one to whom she spoke.	_____	_____
c) The boy smiled and then she spoke to him.	_____	_____

9. At the winter carnival, each event was organized by a teacher and a student. Tom and Miss Clark, the new teacher, were in charge of skating, which was to follow the snowshoe event. As soon as the snowshoe race was over, the new teacher and Tom wanted to start skating. However, they waited for the prizes to be awarded before they went to the microphone and told the skaters to begin.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) It was skating that the new teacher and Tom wanted to start.	_____	_____
b) The new teacher and Tom were going to skate and they wanted to start.	_____	_____
c) The new teacher wanted Tom to start skating.	_____	_____

10. Most people wear clean socks and shoes, but not Larry. Every day he plays baseball. By the end of the week, his running shoes are covered with grass stains and dried mud. On Saturday morning, Larry puts on clean clothes, clean socks, and the same dirty shoes. Then he rushes out of the house before his mother has a chance to throw his runners into the washing machine.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Most people wear clean socks and clean shoes.	_____	_____
b) Most people wear shoes and clean socks.	_____	_____
c) Most people wear socks and clean shoes.	_____	_____

11. As usual, Janet invited Marie to sleep at her house on Friday night. They always ate supper with the family, helped with the dishes, looked at magazines in Janet's room, and then watched the late movie together. This Friday they had an argument, and Janet's mother heard the front door slam at eleven-thirty. Walking slowly into the room, Janet explained that her best friend had left early tonight.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Tonight her friend had best leave early.	_____	_____
b) Early tonight was when her best friend had left.	_____	_____
c) Tonight her best friend had left early.	_____	_____

12. Bobby and Tom were sitting in the backyard when they heard Ruff barking as he came around the house. Quickly they each found a hiding place. Bobby crawled under the lilac bush, and Tom dodged behind the outdoor fireplace. The dog found Tom more quickly than Bobby. He probably saw the boy's shoes beside the wood pile.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) It was Bobby who was found by the dog more quickly than Tom.	_____	_____
b) The dog found Tom more quickly than the dog found Bobby.	_____	_____
c) The dog found Tom more quickly than Bobby found Tom.	_____	_____

13. Tony was the art teacher's helper today. He passed out paper and brushes. He mixed paints. And when he ran out of powdered red paint, he went to fetch the red paint box from the storeroom. He had to read the labels, because all the paint boxes looked the same. When he found the right box, he hurried back to the room to mix the rest of the paint.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) He went to fetch the box of red paint.	_____	_____
b) It was the red box of paint that he went to fetch.	_____	_____
c) He went to fetch the red paint for the box.	_____	_____

14. The party had been going on for two hours when Harry took over the microphone and began playing his guitar. He sounded terrible. Everyone stopped dancing and started talking. Harry took the hint and left the stage. When the band began playing again, the young people stopped talking and began dancing once more. The dancing teenagers knew how good music sounded. They liked the band's music, but not Harry's.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) The dancing teenagers knew how to sound good at music.	_____	_____
b) The dancing teenagers knew how good the sound of music was.	_____	_____
c) The teenagers who were dancing knew how music which is good sounded.	_____	_____

15. After school the teachers had to meet with the principal to talk about the timetable for next year. As the meeting went on, the air in the room became very stale because some of the teachers were smoking. Finally the principal asked the teachers to stop smoking and opened the windows. Soon the room was full of fresh air again and the meeting continued.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) The principal asked the teachers to stop others from smoking.	_____	_____
b) Other people were smoking and the teachers asked the principal to stop them.	_____	_____
c) The teachers were smoking and the principal asked them to stop.	_____	_____

16. Most people who shop like pleasing salesmen. For that reason, some salesmen take classes which help them learn to be courteous and friendly, even with customers who are not. This training pays off for the businesses which employ them, since many customers walk out without buying anything if they are unhappy with the salesman.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Salesmen who are pleasing are liked by most people who shop.	_____	_____
b) Most people who shop like to please salesmen.	_____	_____
c) Salesmen like pleasing most people who shop.	_____	_____

17. Michael's family was very proud that his baseball team was in the national finals. His father had flown with him to Toronto to watch the game in person. His mother and sister had gone to his grandmother's house to watch it on television with her. And Michael's brother watched the game on the patio where he had taken the portable television. They all saw him hit the winning home run.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) The game on the patio was watched by Michael's brother.	_____	_____
b) While on the patio Michael watched his brother's game.	_____	_____
c) While on the patio Michael's brother watched the game.	_____	_____

18. The mailman kept biscuits and candies in his pocket to munch on while he was making his rounds. When he delivered mail to the Martin house that day, he gave Mary Ellen some of the candies. Butch, her dog, wanted something to eat, too. But he threw her dog biscuits rather than candies, because they were easier to chew.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) It was dog biscuits that he threw her rather than candies.	_____	_____
b) He threw her candies rather than dog biscuits.	_____	_____
c) He threw biscuits rather than candies to her dog.	_____	_____

19. Mr. Smith's drug store was across the street from a school. Every afternoon, friendly groups of boys and girls crowded into the store to buy treats. Most of the boys bought ice cream bars with their money. The girls, however, bought licorice, candy bars, or potato chips. Mr. Smith decided that boys like ice cream better than girls.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) Boys like ice cream better than girls like ice cream.	_____	_____
b) Boys like ice cream better than girls like boys.	_____	_____
c) It is ice cream that boys like better than they like girls.	_____	_____

20. Central School's field hockey team had a fund-raising drive for money to buy uniforms. Three team members submitted designs for the uniforms, and a vote was taken at a general meeting. <u>The team's selection was announced in the morning</u>. They had chosen blue shorts, blue-and-gold shirts, and gold knee socks.

	GIVES A MEANING	DOES NOT GIVE A MEANING
a) The morning's selection was announced by the team.	_____	_____
b) That the team was selected was announced in the morning.	_____	_____
c) The selection made by the team was announced in the morning.	_____	_____

APPENDIX F

ITEM ANALYSIS OF (CAT)

ITEM ANALYSIS OF CAT

Test Item No.	Biserial Correlation	Item Difficulty Index
1.	.440	.198
2.	.617	.171
3.	.685	.319
4.	.535	.260
5.	.779	.330
6.	.538	.254
7.	.325	.134
8.	.546	.136
9.	.766	.377
10.	.832	.343
11.	.677	.315
12.	.679	.306
13.	.647	.314
14.	.668	.295
15.	.612	.153
16.	.477	.235
17.	.640	.309
18.	.631	.315
19.	.526	.188
20.	.767	.351

APPENDIX G

ITEMS READMINISTERED FRO[illegible]ND CAT

APPENDIX G

ITEMS RE-ADMINISTERED

			SIT			CAT	
Student	Grade	RS	UA	SS	US	SS	US
A	7	High	5,38	12,30*	14,25	11,14	9,16
B	7	High	29,31	15,30	14,25	14,17	11,16
C	7	Low *	11,31	17,30	14,25	11,14	3,9
D	7	Low *	5,38	15,30	1,28	13,17	16,19
E	7	Low *	31,38	17,30	14,25	14,17	9,16
F	7	High *	38,5	17,30	14,25	11,14	3,20
G	6	[redacted] *	5,38	12,15	14,25	11,14	9,16
H	6	High *	29,38	15,30	2,25	14,17	3,20
I	6	High	5,31	15,30	1,14	4,13	16
J	6	Low	31,38	12,15	14,36	13,14	9,16
K	6	High *	11,38	9,30	14,25	14,17	-
L	6	Low	5,38	15,30	14,25	14,17	7,16

* designates students to whom the CAT was administered before the SIT in the original test administration.

RS = Reader Classification
UA = Unambiguous Sentences
SS = Surface Structure
US = Underlying Structure