

11132

NATIONAL LIBRARY
OTTAWA



BIBLIOTHÈQUE NATIONALE
OTTAWA

NAME OF AUTHOR..... Preston Donald Giffen

TITLE OF THESIS..... The Relative Difficulty of Selected Syntactic
Structures of English as Measured by a Cloze
Test.

UNIVERSITY..... University of Alberta, Edmonton

DEGREE FOR WHICH THESIS WAS PRESENTED..... Ph. D.

YEAR THIS DEGREE GRANTED..... 1972

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.

(Signed)..... *A. Donald Giffen*

PERMANENT ADDRESS:

40 University Avenue

Wolfville, Nova Scotia

P. O. Box 1024

.....

DATED..... February 291972

UNIVERSITY OF ALBERTA

THE RELATIVE DIFFICULTY OF SELECTED SYNTACTIC
STRUCTURES OF ENGLISH AS REVEALED
BY A CLOZE TEST

by



PRESTON DONALD GIFFEN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

SPRING, 1972

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "The Relative Difficulty of Selected Syntactic Structures of English as Revealed by a Cloze Test" submitted by Preston Donald Giffen in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

[Signature]
Chairman
W. Hohmann
E. B. Buxton
G. D. Richards
J. H. Hood

Date *Dec. 9, 1971*

ABSTRACT

An instrument employing "Cloze" procedure and incorporating each of 38 transformational structures of English three times was prepared and administered to students of mixed ability in grades seven, nine, and eleven. Types of transformations included embedding, simple, position shift, deletion, and conjunction. The total number of acceptable and unacceptable responses to each structure of the instrument were used to prepare a list of relative difficulty of those structures.

A difficulty index for each transformation was calculated by subtracting the percentage of acceptable responses on the "complete and delete" instrument from unity. Each student participating in the study was assigned a syntactic ability score by crediting him with the difficulty index figure for each acceptable response made.

Each student also wrote two short essays which were marked by three markers using the general impression method. Papers were placed in one of five equal piles and assigned values from one to five. Scores for two papers marked by three markers were totalled for each student.

Students' syntactic ability scores were correlated with their cumulative writing scores using the Pearson Product Moment method. A correlation of .73, significant at the .01 level, was obtained.

Except for deletion transformations, subjects showed evidence of developmental patterns in overall syntactic ability across the grade levels of the study significant at the .01 level. Results also clearly showed a similarly significant developmental pattern in writing ability across the grade levels.

Difficulty rankings of particular transformations across grades and between groups were calculated using the Spearman Rank Order correlation method. Rankings correlated beyond the .01 level of significance for embedding transformations, and beyond the .05 level for simple transformations. Only the relationship between rankings in grades nine and eleven reached the .05 level of significance for position shift transformations.

Further findings supported earlier research indicating that syntactic structures common to speech are relatively easier to use than those common to writing. Structures in a position common in speech were found to be easier than the same structures used in positions common only to writing. Possible reasons for this finding were suggested.

The major weakness of the study involved use of deletion test items, which did not seem capable of any clear differentiation across grades.

Suggestions for future research included similar studies limiting items to those capable of clearer distinction of more capable writers (for example, embedding transformations).

ACKNOWLEDGEMENTS

The investigator is grateful to the Department of Secondary Education, Faculty of Education, University of Alberta, and to the Province of Alberta Scholarship Fund for financial assistance which made this study possible.

The Edmonton Public School Board, and teachers and students of D. S. MacKenzie Junior High School, Ritchie Junior High School, and Strathcona Composite High School are thanked for their co-operation in making collection of the sample possible.

Special thanks are owed to Dr. R. G. Martin, Committee Chairman, for his incisive insights, helpful criticism and guidance throughout.

To Dr. E. W. Buxton and Dr. G. D. Prideaux, the investigator owes thanks for ideas, inspiration and encouragement.

To Sam Robinson, John Oster, and Imelda Forsyth, markers and proof readers, sincere thanks.

To Mrs. Isobel Giffen, whose faith never faltered, deepest affection.

And lastly a dedication to the investigator's wife, Naomi, without whose support and encouragement the first steps would not have been taken, nor the bridge crossed.

TABLE OF CONTENTS

| Chapter | Page |
|--|------|
| I. THE PROBLEM | 1 |
| Background | 1 |
| Statement of the Problem | 6 |
| Limitations | 8 |
| Significance of the Study | 10 |
| Definitions of Terms | 12 |
| General | 12 |
| Specific Definitions | 16 |
| Overview and Organization of the Study | 23 |
| II. REVIEW OF THE LITERATURE | 25 |
| Reading Difficulty, Comprehension and "Cloze" Procedure | 26 |
| Introduction | 26 |
| Early Readability Formulas | 29 |
| More Detailed Formulas | 29 |
| More Efficient Approaches | 30 |
| The "Cloze" Technique | 33 |
| Summary | 42 |
| Grammars and Linguistics | 44 |
| Summary | 71 |
| Psycholinguistics and Transformational Theory. | 74 |
| Introduction | 74 |
| Segmentation | 80 |
| Syntax, Semantics, and Comprehension | 83 |

| Chapter | Page |
|--|------|
| Deep and Surface Structure | 91 |
| Transformation | 93 |
| Summary | 108 |
| Segmentation | 108 |
| Syntax, Semantics, and Comprehension . . . | 109 |
| Deep and Surface Structure | 109 |
| Transformation | 110 |
| General Summary of Psycholinguistics . . . | 111 |
| Syntactic and Writing Ability | 113 |
| Introduction | 113 |
| Language Ability and Development | 114 |
| Summary | 131 |
| Writing Applications | 133 |
| General Summary | 141 |
| The Evaluation of Writing | 142 |
| Summary | 146 |
| III. THE GRAMMAR OF THE STUDY | 147 |
| Glossary of Grammar | 150 |
| The Transformational Component | 154 |
| Transformations of Cloze Test | 155 |
| Transformations of the Grammar | 156 |
| Transformations | 157 |
| IV. THE TEST INSTRUMENT | 173 |
| General Design | 173 |
| Pilot Study | 176 |
| Administration | 178 |

| Chapter | Page |
|--|------|
| Scoring | 179 |
| V. THE WRITING ASSIGNMENTS | 182 |
| Physical Surroundings | 182 |
| Time and Length of Assignments | 183 |
| The Sample Size | 183 |
| Topics | 184 |
| Summary | 186 |
| VI. DESIGN OF THE STUDY AND PROCEDURES | 187 |
| Introduction | 187 |
| The Subjects and the Sample | 190 |
| Test Design | 193 |
| Test Administration, Writing Sample Collection, and Marking | 198 |
| Complete and Delete Test | 199 |
| Treatment of Data and Statistical Analysis to be Applied | 203 |
| Significance Level | 205 |
| VII. FINDINGS OF THE STUDY | 207 |
| "Complete and Delete" and Transformation Difficulty | 207 |
| Mean Scores on "Complete and Delete" Test | 207 |
| Transformation Difficulty as Indicated From Test Scores | 208 |
| Transformation Difficulty Within Groups | 214 |
| The Writing Assignment and Inter-Marker Reliability | 215 |
| Mean Scores on Writing Tests | 215 |
| Inter-Marker Reliability | 217 |

| Chapter | Page |
|--|------|
| Principal Correlation for the Group as a Whole | 218 |
| Interrelationship of "Cloze" and Writing Score With I.Q. Factored Out | 220 |
| Correlation Matrices for All Factors by Grades | 222 |
| Additional Findings for Grade Seven | 226 |
| Additional Findings for Grade Nine | 227 |
| Additional Findings for Grade Eleven | 227 |
| Additional Findings for Trends Across Grades | 228 |
| Summary of Findings | 229 |
| VIII. RE-EXAMINATION OF HYPOTHESES | 232 |
| IX. SUMMARY, CONCLUSIONS, AND IMPLICATIONS | 236 |
| Summary of the Study | 236 |
| Findings and Conclusions | 240 |
| The Difficulty of Transformational Structures | 240 |
| Developmental Trends in Syntactic Ability | 252 |
| The Relationship of Syntactic Ability and Writing Ability | 255 |
| Summary and Reaction to Hypotheses | 258 |
| Limitations and Applicability of Findings | 260 |
| Suggestions for Further Research | 262 |
| Implications of the Study | 263 |
| Concluding Statement | 266 |
| BIBLIOGRAPHY | 268 |
| APPENDICES | 284 |

LIST OF TABLES

| Table | Page |
|---|------|
| 1. Lorge-Thorndike, Form 4, Intelligence Quotient Data (Verbal) | 191 |
| 2. Number of Subjects by Grade and Sex | 192 |
| 3. Qualifications and Experience of Markers of The Study | 202 |
| 4. Means of Scores on "Complete and Delete" Test | 207 |
| 5. Difficulty of the Transformations of the "Complete and Delete" Test | 209 |
| 6. Relative Difficulty of Embedding Transformations by Grades | 211 |
| 7. Relative Difficulty of Simple Transformations by Grades | 212 |
| 8. Relative Difficulty of Position-Shift Transformations by Grades | 213 |
| 9. Relative Difficulty of Deletion Transformations by Grades | 213 |
| 10. Means of Scores on Writing Assignments | 216 |
| 11. Pearson Product-Moment Correlation Matrix for Six Factors of Test Population | 221 |
| 12. T Values and Probabilities Associated with them for Above Correlation Matrix | 221 |
| 13. Pearson Product-Moment Correlation Matrix for Six Factors for Grade Seven | 223 |
| 14. T Values and Probabilities Associated with them for Above Correlation Matrix | 223 |
| 15. Pearson Product-Moment Correlation Matrix for Six Factors for Grade Nine | 224 |
| 16. T Values and Probabilities Associated with them for Above Correlation Matrix | 224 |

Table

Page

| | | |
|-----|---|-----|
| 17. | Pearson Product-Moment Correlation Matrix for Six Factors for Grade Eleven | 225 |
| 18. | T Values and Probabilities Associated with them for Above Correlation Matrix | 225 |
| 19. | Distribution of High and Low Writing Marks by Grades | 256 |
| 20. | Distribution of High and Low Syntactic Ability Marks by Grades | 257 |

LIST OF FIGURES

| FIGURE | PAGE |
|---|------|
| I. Finite State Model | 52 |
| II. Phrase Marker | 54 |
| III. Chomsky's <u>Aspects</u> Grammar | 62 |
| IV. A Feature Matrix | 63 |
| V. Suggested Speech Production Model | 77 |
| VI. Yngve's Depth Models | 84 |

CHAPTER I

THE PROBLEM

The intimate relation between language and the thought processes makes a study of language development important . . .¹

1. BACKGROUND

A primary task for the teacher of English is to attempt to "improve" the quality of the written and spoken linguistic output of the students. This task is complicated by the difficulty found in evaluating student output (and specifically written output). Experienced teachers of English feel themselves capable of judging the relative merit of student essays or themes using criteria which might be divided into the following headings:

1. Content
2. Accurate usage
3. Correct spelling
4. Adequate punctuation
5. Unity and organization
6. Coherence
7. Extensive vocabulary
8. Stylistics and emphasis

It is simple for the English teacher to be objective about the

¹Lou L. LaBrant, "A Study of Certain Language Developments of Children in Grades Four to Twelve, Inclusive," Genetic Psychology Monographs, XIV, No. 5 (November 1933), p. 393.

accuracy of usage, spelling and punctuation. It is somewhat more difficult to judge the unity, coherence, and organization of a piece of writing. The presence of an extensive vocabulary is also objective. However, it can still become extremely difficult to pin down stylistic qualities which mark the "better" or "more mature" writer. The experienced reader can often say of an essay that it has a pleasing flow, a "facileness," a quality of logical organization and development which serves to emphasize what is important. Few such readers, however, can specify the particular objective features of such "good" writing which set it apart.

One quality often possessed by such "better" or "more mature" writing is what might be termed "compression," a quality of saying much through a minimum of words. This compression, however, seems to go hand in hand with a certain complexity which is established by the use of patterns of logic, contingencies, dependencies, and subordinations--all of which contribute to the "internal coherence" of any passage. While several of these qualities exist beyond the sentence unit, many of them operate within single sentences. In other words, it appears that the appropriate use of a wide variety of language structures of more or less complexity is one important feature which sets off "superior" or "more mature" writing.

Whether the students' mastery of such structures of increasing complexity is related simply to the process of maturation is one aspect examined in the present study,

in which patterns of development are investigated.

The criteria of "complexity of language structures" are to be developed using the terminology and theory of generative-transformational grammar. Thus this study serves to a small extent to examine the psychological reality of one aspect of the generative-transformational theory. As will be made clear further in this investigation however, no claim is made as to the psychological reality of such a theory. Indeed, this theory serves the present study only insofar as it provides a systematic set of labels or descriptions for certain structures of language which this particular grammar is capable of producing. The one aspect of this theory upon which the present study reflects is what has been termed the "derivational theory of complexity," which claims that language structures involving a greater number of steps or operations in their derivational history within the generative-transformational theory will present the greatest difficulty to learn and to use. This particular theory has been both attacked and defended; perhaps this study can throw further light upon its viability.

The test instrument of this study is a "Cloze" test, a type of test used in the past as a measure of readability. In fact, the use of such a test reflects the commonly-held belief that students who can read well are also better writers. This bears on whether the same quality, a "syntactic awareness," is an important factor in both reading and writing ability. The theory underlying the use of the

"Cloze" text is that students capable of accurately filling in words deleted from such a test show that they comprehend the passage in question. The present study extends the above theory by examining whether the student capable of producing certain structures of language by inserting key words deleted from such structures is familiar with, or knows, or comprehends such structures, whether syntactic or semantic in nature.

Note that the procedure above is not a "multiple choice" situation. The student must pull the appropriate word from his own language competence influenced only by the context surrounding the blank confronting him.

The investigator began preparation of the materials for this study with the intent of merely investigating the relative difficulties of various structures of English language. A small generative-transformational grammar was prepared using a skeleton lexicon and sufficient transformations to produce many of the more common structures of English syntax. Then a series of four short passages was prepared involving somewhat differing genres of writing (i.e., narrative, descriptive, expository, etc.) so as to involve most of the more common language structures. Then, key words were deleted from the passages to create what might be called a "structurally constructed 'Cloze' test." Pilot testing and revision resulted in an instrument incorporating three occurrences of each of 38 items.

One technique incorporated in the instrument of this study which does not have the extensive background of the Cloze

test is what is here termed a "deletion" item. The process of production of written language involves not only decisions as to what might be added, but also judgements as to what might be left out. Economy and a lack of redundancy have been considered further indicators of the superior writer. The investigator was not wholly satisfied with the test technique selected, that of underlining a passage containing a word or words which might be deleted without substantially changing passage meaning. It was felt that this type of item was akin to the multiple-choice question wherein it is only a question of selecting from a limited number of choices. Would such items result in disproportionately high scores on such deletion items? Pilot testing indicated a proportion of errors sufficiently large to dispel such fears, and consequently the deletion items were retained.

The next question which arose was of what use the information provided by such an instrument might be to English teachers and persons involved in English curriculum? The test might indicate the relative difficulty of its selected items (information interesting to the linguist, but of little value to the English teacher), but there would be no proof that students capable of dealing accurately with the test items were "superior" or "more mature" writers. Thus it was decided that a writing sample should be collected from each of the participating students, a sample which could be evaluated so that a comparison might be made between "cloze-delete test" scores and writing scores. Exact techniques

used in the production of the above-mentioned test and in the collection of the writing sample will be dealt with in the chapters entitled "Design of the Study" and "The Instrument."

It is hoped that the present investigation might shed some further light on areas of interest to English teachers and curriculum designers, linguists, and psycho-linguists. The exact nature of these areas will be discussed in the section entitled "Significance of the Study" to follow.

2. STATEMENT OF THE PROBLEM

This study is concerned with the relative difficulty of certain selected syntactic structures of English as revealed by a structurally-constructed "cloze and delete" test. It also seeks to investigate the relationship between students' ability to reconstruct certain syntactic structures and their general writing ability. It further seeks to identify developmental patterns of students' syntactic ability over the spans of grades seven, nine, and eleven.

More specifically, the study has attempted:

1. To construct a "cloze and delete" test which indicates, by means of frequency of appropriate response, the relative difficulty of certain syntactic structures of English.

2. To administer the above test to pupils in grades seven, nine, and eleven.
3. To assign two writing topics to each of the students writing the above test, samples to be written in a given time and to be of approximately the same length.
4. To administer the above topics and to collect the writing samples, one before administration of the "cloze and delete" test, and one after.
5. To determine the relative difficulty of the syntactic structures incorporated in the "cloze and delete" test by means of frequency of appropriate response.
6. To compute overall syntactic ability of individual students by adding scores from the student in "cloze and delete" test weighted to indicate relative difficulty.
7. To evaluate wholistically the two writing samples from each student, each of three markers dividing each group of papers into five equal piles of increasing ability or maturity.
8. To examine the correlation between individual student's syntactic ability score and his composite writing score.
9. To examine overall patterns of development in both writing ability and syntactic ability using average scores from each class group.
10. To examine patterns of development in syntactic ability for particular types of syntactic structures (deletions, conjunctions, embeddings, simple transformations) across class groups.

A statement of hypothetical conclusions will further define the problems of this study:

Null Hypotheses

- (a) Individual student's additive syntactic ability scores will not correlate significantly with their individual total writing scores.
- (b) No significant patterns of development in syntactic ability as measured by the instrument of this study will appear across grades seven, nine, and eleven.

- (c) No significant patterns of development in writing ability as determined by the methods of this study will appear across grades seven, nine, and eleven.
- (d) No significant patterns of development in ability with the various types of transformations of this study (embedding, simple, position shift, and deletion) will appear across grades seven, nine, and eleven.

The level of rejection of the above hypotheses is to be a probability of 0.05.

3. LIMITATIONS

This study makes no claim as to the psychological reality of the processes of sentence production and transformation outlined in its limited generative-transformation grammar. Further, no claim is made that the grammar of this study is complete or comprehensive. It is merely adequate to produce the syntactic structures which are incorporated in the "cloze and delete" test instrument which this study utilizes.

The "cloze and delete" test does not incorporate all possible syntactic structures, but only a number of common ones. Also, each such structure occurs three times, thus severely limiting, for research purposes, the wide variety of natural contexts in which it may occur.

While every effort was made in constructing the "cloze and delete" test to avoid including passages and sentences which might cause reading difficulty, and while several students whose school records indicated reading difficulties were eliminated from the sample, it is possible that

some responses classified as "inappropriate" were the result of reading or vocabulary difficulty. What steps were taken to avoid this problem are outlined under "Procedures."

The marking method used by the three markers on each of the two sets of papers was a "wholistic" one. That is, decisions as to the "quality" of a piece of writing were made in a time period from two to three minutes on a "general impression" basis. Thus errors in mechanics received less attention than they might have had a more thorough marking scale been employed. Consequently, some objectivity has been sacrificed in order to emphasize the student's "overall language facility;" that is, his ability to use language as an effective means of communication of ideas. The specific instructions given to the markers are outlined in the chapter on "Design of the Study and Procedures."

Time limitations necessitated that the sample essays written by the students be first drafts rather than revised drafts. Consequently, the writing samples tend to be somewhat lacking in "polish." However, the investigator felt that this limitation might not in truth be one. The style of writing which students would utilize on a first draft might be closer to their actual thought patterns than any revised drafts which might have been produced.

4. SIGNIFICANCE OF THE STUDY

This study attempts to bring together techniques from the fields of reading and linguistics in order to explore whether certain syntactic structures of English are more difficult for students to reconstruct than others, and whether students of varying intelligence and age level gradually increase in their ability to recognize or respond to structures of greater syntactic complexity or difficulty. A further aim of this study is to examine whether it is possible by means of a structurally-created "Cloze" test to make objective evaluations of students' writing ability. A positive result in this inquiry would tend to support the theory that there is a relationship between students' ability to read and their ability to write, as a "Cloze" test would seem to be a good test of both readability and reading ability. Further, should a strong positive correlation be found to exist between the measure of syntactic ability of individual students and their scores obtained from writing samples, this study would provide a powerful diagnostic tool for the teacher of composition, and might also indicate directions which the teaching of composition might take.

Even in the absence of strong positive correlations, the results of this study might give some indication of the importance of the syntactic component in judging the relative quality of students' writing. The present situation, in which only very limited information as to discriminating factors between poorer and superior writers has been developed, has

not contributed much to improvement of the pedagogy of teachers of composition. Any contribution which this study might make would further the research needed in this vital area. Of particular interest might be some indication as to whether there is some kind of systematic progression of complexity of structures which students use with increasing maturity. It is hoped that this study might provide precisely such information.

In the area of linguistics, this study provides additional information as to the viability of the derivational theory of complexity. Examination of the results indicating "difficulty" might provide an empirical starting point for investigations into the nature of linguistic performance. At present, there is all too little empirical information of this type.

A further contribution which this study might make is to support the extension of the uses of a "Cloze" type test into the area of syntax suggested by Fagan.² It seems to the investigator that such a test makes demands upon the subject's sensitivity to context, and particularly to structural context.

The concepts of linguistic "competence" and "performance" are difficult to place on an empirical basis, but not too difficult to define. "Competence" represents the native speaker's knowledge of his language, and

²William J. Fagan, "An Investigation into the Relationship Between Reading Difficulty and the Number and Types of Sentence Transformations." (Unpublished Doctoral dissertation, University of Alberta, 1969).

performance is what, in fact, he actually does. The test which is the principal instrument of this study seems to involve two distinct operations: recognition and production. The subjects must first recognize or demonstrate familiarity with the structures of the test. Such a process calls upon knowledge or sensitivity which in many cases (in the investigator's view) goes beyond the sphere of the subject's everyday world of language performance. The subject's ability to fill in or cross out correctly to complete the test is an act of performance, but not purely so. In fact, it might very well be true that in the case of some subjects, the level of competence might be so much higher than the level of performance during the creative process that the correlation of test results against writing performance might possibly be negative. Such a result, while running counter to naive expectations, would be of great interest to both linguists and teachers.

5. DEFINITIONS OF TERMS

A. General

This portion of the study discusses in general fashion some of the concepts and terminology which are necessary to the investigation. A section of more specific definitions will follow.

A grammar is a theory of a language--a set of precepts or rules, a list of operations, and a lexicon. A grammar produces "strings" of linguistic output which the native

speaker of the language will acknowledge to be acceptable strings of the language in question. While a grammar may incorporate optional elements or operations, once the available choices have been made its operation is completely explicit. Thus a complete grammar of a language would specify all of the acceptable (and only the acceptable) sentences which constitute that language.

The various sub-branches of the various types of grammars available are named functionally or, in some cases, historically. The most interesting aspect of the fact of the existence of so many competing grammars of, for instance, English, is that they are all dealing with the same corpus, a body of linguistic data which developed in accordance with no known specific theory, but rather organically. The fact that linguists are prepared to attempt to formulate "rules" for such a collection of data (no matter how successful they are) says much for the belief of most students of behavior that some moving force or cause underlies all change. Thus the linguist shares many of the problems of experimental psychologists, although at least the linguist's data is overt.

Few investigations into the reading aspect of language studies have attempted much penetration into the complexities of sentences beyond acknowledging that the longer a sentence becomes, the more difficult it becomes to comprehend. Most tests of "readability" incorporate the concept of "frequency of hard words," essentially a dependency upon vocabulary

factors. Tests of subject comprehension make demands upon short-term retention of facts and concepts by subjects, and make little or no reference to the organization of such material. Rawson³ used variation in patterns of logic to determine the effect of such variation on subjects' comprehension, and found by factor analysis that such variations accounted for approximately 20% of observed variation. Fagan⁴ investigated the matter of the effect of various syntactic patterns on comprehension using a method similar to the present study and noted considerable variation in the effect of occurrence of certain structures on comprehension. Perhaps the present study will serve to further this investigation and help to extend the concept of "readability."

The problem of defining a "mature writing style" is considerable. However, most critics of written language would agree with the following: as writers mature, their output becomes more succinct, more efficient, more cohesive, and less redundant. The better writer's style is reflected by an ability to put ideas and concepts together, and to make them flow progressively by means of devices of liaison, subordination, and conjunction. The mature writer also seems conscious of having an equally mature audience or reader whose intelligence he does not insult. However, the better writer

³Hildred I. Rawson, "A Study of The Relationship and Development of Reading and Cognition," (Unpublished Doctoral dissertation, University of Alberta, 1969).

⁴Fagan, op. cit.

is also capable of adjusting his style to the task at hand; he is flexible and adaptable, yet consistent to purpose. For the mature writer, writing becomes an efficient tool, marvelously adequate to its user's purpose, yet never drawing attention to itself.

General impression or "wholistic" evaluation of written themes is a method which would seem to invite a lack of agreement of readers. This method, whereby evaluators using fairly general criteria spend no more than two or three minutes on a theme, seems particularly vulnerable to the charge that its subjectivity destroys any reliability claimed for it. Even should high degrees of inter-rater reliability be obtained, such ratings might not even be the result of the readers reacting to similar criteria or features of the writing sample. As outlined in Chapter 2, Section 5 of this study, the alternative, analytic marking, can achieve great reliability at the cost of destroying the effect of certain stylistic factors. What can be said for the general impression method of marking and its application to this study is that it invites reader reaction to the grosser structures of the writing sample and to the flow of ideas conveyed by them. As these gross structures constitute the syntactic units of the writing specimens, it would appear on the surface that general impression marking would be particularly acceptable for this study provided that the markers were not directed specifically to be aware of

complexity of sentence structures. The specific instructions to the markers of the study are set out under "Procedures."

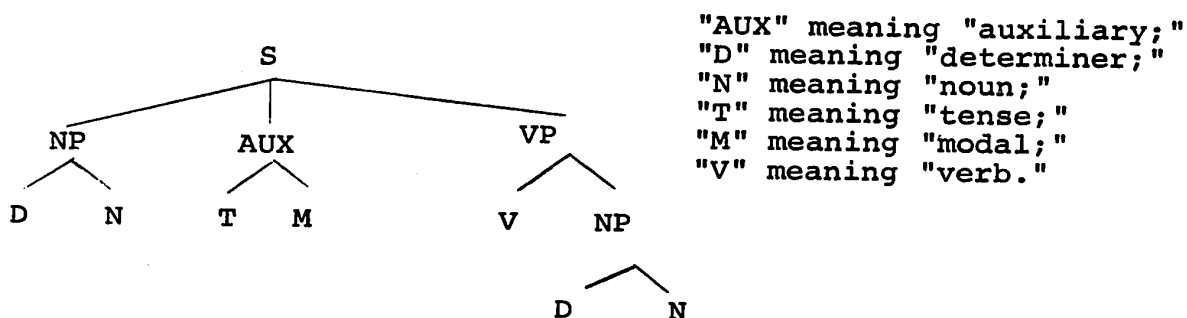
B. Specific Definitions

1. Grammar: a theory of a language incorporating a list of morphemes, non-terminal categories (S, NP, VP, etc.), a set of obligatory and optional operations, and a set of rules for the combination of elements into expressions of that language.
2. Morpheme: the minimum meaningful linguistic unit.
3. Phoneme: the minimum structural unit in the sound system of a language.
4. Syntax: the descriptive analysis of linguistic forms of a language larger than words; the manner in which the morphemes of a language may be combined to produce the functional units of that language.
5. Semantics: the "meaning" system of a language whereby morphemes are said to convey denotative and connotative notions; the study of the "meaning" of words.
6. Phonology: the study of sounds (in the case of this investigation, of the sound system of a particular language).
7. Lexicon: the list of morphemes which constitute a language; one major element of a grammar of a language.
8. Rewriting rule is of the general form $A \rightarrow B$, where A is a single nonnull symbol and B is a nonnull string of symbols distinct from A.

9. Derivation involves sequential application of rewriting rules beginning with an initial symbol S (meaning "sentence") and continuing with each step involving the application of one rule to one symbol in a preceding string.

10. Phrase structure rules: A series of ordered rules specifying the obligatory and optional operations by means of which a symbol (such as "S," meaning "sentence") is to be rewritten as a number of other elements; e.g. $S \rightarrow NP+VP$, "S" meaning "sentence," NP meaning "noun phrase," VP meaning "verb phrase," the symbol \rightarrow meaning "is to be rewritten as"

11. Phrase marker (or P-marker): A "branching tree" diagram which shows by means of symbolically labelled "nodes" (or points) and connecting lines the operation of phrase structure rules.



A sentence is a structured string of words produced by the rules of a grammar. Thus the orthographic representation of a sentence is a structured string of morphemes extending from a capital letter to some form of terminal punctuation.

12. Recursiveness is a property of G.-T. grammars whereby an optional symbol permits the incorporation of one

sentence within another as often as the option is exercised.

e.g. $S \rightarrow NP + AUX + VP \dots$

$NP \rightarrow (D) N (S)$

13. Embedding is the process of exercising the above option so that one sentential element is incorporated within the phrase marker and phrase structure of another "matrix" sentence.

14. Transformation is an obligatory or optional operation whereby a phrase structure meeting specified Boolean conditions on analyzability (or a structural description) is changed into different structure by any or all of the following processes: substitution, addition, deletion.

15. Transformational history refers to the number of transformational rules that have been applied during the derivation of a sentence.

16. Traditional grammar is a type of grammar which seeks to describe the data of a language (and occasionally, to provide a history of such changes as have occurred). Such grammars, if they approach completeness, are of great complexity and length. e.g. those of Poutsma,⁵ Kruisinga,⁶ Jespersen,⁷

⁵Hendrik Poutsma, A Grammar of Late Modern English, second ed. Groningen: P. Noordhoff, 1914-1929.

⁶Etsko, Kruisinga, A Handbook of Present-Day English, fourth ed. Utrecht: Kemink en Zoon, 1925.

⁷Jens Otto Harry Jespersen, A Modern English Grammar on Historical Principles. Copenhagen: Ejnar Munksgaard, 1909-1949.

and Zardvoort.⁸

17. Prescriptive grammar: (also called pedagogical or school grammar) is the product of a tradition of linguistic thought which seeks to specify "correct" usage within a language. Such grammars often used Latin grammar as a model for English. e.g. Lowth, Murray.

18. Structural grammar defines the units of language structurally (or positionally), morphologically, and functionally. It treats languages as unique entities, and considers speech (rather than writing) to be primary. Structural grammarians identify basic "sentence patterns" such as S - V - O (subject, verb, object) or N - V_t - N (noun, transitive verb, noun).

19. A generative-transformational grammar consists of a set of base phrase structure rules, a lexicon, a set of transformational rules, a phonological component, and a semantic component. While a G.T. grammar may contain options, it is entirely explicit in its operation.

20. Deep structure is the underlying structure of a sentence, the structured string produced by the phrase structure or "P.S." rules and the first lexical insertion of a G.T. grammar. Early G.T. theory stated that deep structure carried the entire meaning or semantic load of the sentence.

2. Surface structure is the final spoken or written form of a sentence. It is produced by the operation of the

⁸Reinard Willem Zardvoort, A Handbook of English Grammar. London: Longmans, 1957.

transformational component of a G.T. grammar upon the deep structure.

22. Generative semantics is a theory of transformational grammar which rejects the concept of a specific level of deep structure containing lexical items. This theory would claim, for instance, that both "The dog was killed" and "The dog was caused to be dead" would come from one phrase marker which would be in itself the semantic representation of both, but which could have either realization.

23. Case grammar is another alternative transformational grammar which has a phrase structure system incorporating a node "Proposition" under which different "cases" of nouns (genitive, dative, accusative, etc.) appear in a certain order. While case grammar can solve some of the difficulties of ordinary G.T. grammar, it has problems of its own, being highly semantically dependent, as well as having highly abstract appearance.

24. Competence is a term used in G.T. grammatical theory to refer to the total linguistic potential or capability of an idealized native speaker of a language.

25. Performance is a G.T. grammatical term which describes what a speaker of a language actually produces in the way of linguistic output, whether verbal or written.

26. Number of transformations refers to the total transformational rules utilized in the derivation of a sentence.

27. Type of transformation refers in this study to

written surface structures which are the result of the application of one particular transformational rule or class of operations of similar type to the deep structure of sentences, e.g. embedding transformation, position shift transformation, etc. Any one type of transformation will apply to particular structural descriptions or arrangements of categories (or elements) to perform specific structural changes resulting in identifiable surface structures.

28. Derivation Theory of Complexity (D.T.C.) postulates that structures requiring a greater number of transformational processes or steps in their derivation will be "more difficult" to produce, difficulty being measured (in most experiments using this metric) by the amount of time required to perform the experimental derivation. (See Chapter II for further discussion). Studies into this metric do not clearly define what constitutes a transformational step or operation, however. For instance, is a step a complete transformation, or is it one of the operations of addition, deletion, transposition, etc., which make up transformation? For example, is a "passive" transposition of elements one process, or do we count one step each for (1) transposition of elements, and (2) insertion of "was" and "by"? Is each of these insertions a single step, giving a total of three?

29. "Cloze" technique normally involves the deletion of every nth word from a passage of writing. The subject is then asked to "fill in" the missing words. Alternative forms are prepared until every word has, in one form or another,

been deleted. The difficulty of a particular passage is determined by the average number of errors over all of the test forms. Scoring may be on the basis of acceptance of either filling in the exact word deleted only, or supplying a synonym.

30. A structurally-constructed "Cloze" test as used in this study differs from normal cloze tests in that words essential to particular transformational grammatical surface structures are deleted, instead of every n^{th} word. Scoring is on the basis of acceptance of any word which results in the re-creation of the particular syntactic structure involved. Thus a semantically acceptable response which did not re-create the structure desired will not be accepted.

31. Delete items as incorporated in the test instrument of this study consist of underlined parts of sentences from which students are asked to "cross out" as many words as they can without changing the overall meaning of the sentence. Scoring is on the basis of whether the crossing-out operations parallel the transformational deletion process available.

32. Difficult test items are defined for this study as test instrument items having a low frequency of acceptable responses. Thus an index of correct responses ranging from a high of 1.000 to a low of 0.000 would result.

33. An additive difficulty index for test items will be obtained by subtracting "frequency of accurate response" index figures from unity. Thus a test item having a frequency index of 0.666 would have a difficulty index of 0.334. A

subject responding acceptably to two out of three items having such an index would receive an additive score of 0.668 out of a possible score of 1.002.

34. Syntactic maturity, for the purposes of this study, is indicated by the cumulative additive score of difficulty indices the subject obtains on the 132 "Cloze" and "delete" items of the test instrument.

35. General impression evaluation of writing (wholistic evaluation) is a technique whereby evaluators read essays rapidly and judge the writer's relative ability with language on criteria of fluency, logical development, and overall ability to use English as a means of effective communication. For the purposes of this study, mechanical errors in spelling and punctuation are to be ignored.

6. OVERVIEW AND ORGANIZATION OF THE STUDY

This study consisted of three phases:

1. The writing and preparation of a test instrument consisting of four passages incorporating narration, description, and exposition from which certain words were deleted. The sentences of the passages were analyzed by means of the G.T. grammar of the study, and certain words within 38 transformational structures were chosen to be deleted.

2. The administration of the test instrument to two classes each of grades seven, nine, and eleven selected from three schools in the City of Edmonton.

3. The collection of two writing samples on assigned topics from each student writing the test instrument. Topics were varied to attempt to obtain writing involving narration, description, and exposition-argument.

Chapter II reviews the theory and framework under which this study was conducted, including research studies and relevant writings in the areas of language, psycholinguistics, reading, and writing evaluation.

The third chapter outlines the grammar of this investigation and the particular grammatical structures chosen for inclusion in the instrument.

The fourth chapter describes the design of the test instrument, its administration, pilot testing, and revision.

Chapter V deals with the writing assignment.

Chapter VI outlines the experimental design of the study, giving information about the sample, administration and collection of materials, and treatment of data.

The seventh chapter reports the results of the experiment and the statistical analyses of the data.

In Chapter VIII the hypotheses of the study are re-examined in the light of experimental findings.

The final chapter contains the summary, conclusions, and implications.

CHAPTER II

REVIEW OF LITERATURE

This chapter examines research studies and related literature in the areas of (1) reading comprehension and its measurement, (2) linguistics, grammars, and psycholinguistics, (3) growth in language ability, and (4) evaluation of writing.

The first section dealing with reading comprehension and attempts to measure it provides background for the "cloze" technique utilized in this study.

The portion discussing linguistics, grammars, and psycholinguistics serves two purposes: it seeks to justify the selection of a generative-transformational model of grammar for this study rather than various alternatives, and it surveys the various attempts which have been made to explore the psychological realities which lie behind human language capability and production. This latter portion concerns itself particularly with the ways in which humans "encode" or store linguistic information, and the reality of the transformational process.

Studies which have attempted to discover the distinctive areas of growth in language ability which indicate development (or greater maturity) are examined in the third portion of this chapter.

The last section of this chapter examines approaches to the evaluation of student writing in order to justify the

selection of the "general impression" approach used in this study.

1. READING DIFFICULTY, COMPREHENSION,
AND THE "CLOZE" PROCEDURE

Introduction

George Miller, describing what occurs when someone hears a sentence, proposes a sequence of happenings which involves six processes:

1. hearing--the perception of sounds,
2. matching--the process of placing a phonological interpretation upon what is heard,
3. accepting--imposing a grammatical interpretation upon the utterance, and either accepting it or rejecting it as a part of his language,
4. interpreting--imposing a semantic interpretation upon the utterance,
5. understanding--grasping the contextual significance of the utterance, for example, the speaker's intention, and
6. believing--making judgments as to the truth value of the utterance.¹

As this study deals with written material, the first two steps would have to be altered to read "perceiving" rather than "hearing," and "morphological" rather than "phonological." However, the remaining processes would all make up what reading theorists have described as "comprehension." To say that a person comprehends what he has read is, in layman's terms, to say that he "understands" what has just passed before his eyes: he should be able to answer

¹George Miller, "The Psycholinguists," Encounter, Vol. 23, No. 1 (July, 1964), pp. 29-37.

questions concerning what he has read, to make judgements as to grammaticalness, truth value, and meaning within context.

What factors can affect the process of comprehension? This question is vital to many people. Certainly educators want their students to "learn to comprehend," if such a skill is in fact open to pedagogy. Publishers want to be able to gauge the relative difficulty of the books they publish in order to suit them to potential readers. Reading theorists, above all others, know the complexity of the reading process, and the seemingly endless list of factors which seem to complicate it. But beyond purely physical matters such as type-size and layout, what can make a book more readable, more comprehensible? In seeking the answer to this question, we might look at words, at groups of words, at sentences, and at entire passages and their subject matter.

There can be no doubt that subject matter is an important factor in comprehension. Some of the concepts of nuclear physics and higher mathematics require levels of abstract thinking beyond most people's ability, and knowledge of highly technical, specialized vocabularies.

The manner in which sentences are arranged provides a flow of thought, providing contextual clues which aid the reader. Logical pattern of reasoning--inductive or deductive--can make comprehension easier.

Sentences--their length and the intricacies of their structures--can obviously be a major factor in making material easy or hard to understand. One only has to

encounter some of the page-long sentences of Ruskin or Faulkner to realize that choice of sentence structure is a major factor in comprehension. (And in style, as witness a succession of "Dick and Jane" sentences.)

Within sentences, the use of noun, adjectival, and adverbial clauses, and various phrases, can have a considerable effect upon ease of comprehension. For example, a long, highly modified introductory noun clause often requires even an experienced reader to "look back." Sentences can also be right-branching (The man who kicked the dog that chased the cat that killed the rat is dead) or left-branching (A considerably less than well-disciplined child spilt the milk) or "self-embedded" (The man the boy the dog the car hit bit saw fell). While right-branching structures seem quite acceptable, left-branching sentences seem somewhat more awkward, and self-embeddings or "nestings" which involve more than one embedding are beyond most readers' comprehension.

Lexical ability--a knowledge not only of the meaning of words, but of alternate meanings, and of connotative qualities--is the most accessible level of a student's ability to comprehend. The greater a person's vocabulary, the greater the number of ideas or notions he has to form concepts, and the greater the number of concepts he can begin to analyze. Without some knowledge of the meaning of a word, a person may perceive it, and may even believe it to be "grammatical," but he cannot interpret, nor understand, nor form judgments about it.

Early Readability Formulas

Early attempts at determining the "readability" or comprehensibility of printed material used approaches involving the counting of elements: words per sentence, syllables within words, frequency of "difficult" words, and so forth. The only approach made toward judging the difficulty of sentence structures followed the reasoning that more complex sentences would tend to be longer, hence the use of total sentence length.

The first true readability formula, by Lively and Pressey,² used the criterion of whether words appeared on the Thorndike list of commonly used words.³ Such a dependence on vocabulary matters was common in most of the earlier studies, such as that of Vogel and Washburne,⁴ which was suitable for grades 3 to 9, and that of Dolch,⁵ suitable for primer to grade 4 level.

More Detailed Formulas

A greater number of factors became involved in trying to assess material difficulty in the work of researchers such

²B. A. Lively and S. L. Pressey, "A Method for Measuring the 'Vocabulary Burden' of Textbooks," Educational Administration and Supervisor, IX (October, 1923), pp. 389-98.

³E. L. Thorndike, The Teacher's Word Book, New York: Teacher's College, Columbia University, 1921.

⁴M. Vogel and C. Washburne, "An Objective Method of Determining Grade Placement of Children's Reading Material," Elementary School Journal, XXVIII (January, 1928), pp. 373-81.

⁵E. W. Dolch, "Vocabulary Burden," Journal of Educational Research, XVII (March, 1928), pp. 170-83.

as McClusky⁶ and Gray and Leary.⁷ McClusky looked at the number of letters per word, the number of words per sentence, and the number of various kinds of nouns used. Gray and Leary examined a great number of factors, finally selecting only five as significant. Their formula examines: the number of different hard words (not appearing on the Dale List of 769 words),⁸ number of personal pronouns, average number of words per sentence, percentage of different words, and number of prepositional phrases. Applicability was from grade 2 to college. This latter study helped to direct the thoughts of investigators toward the greater complexities of the reading process, but the greatest change in studies was toward greater efficiency.

More Efficient Approaches

Washburne and Morphett's⁹ formula, for grades 1 to 9, looked at the number of different words, the number of words not in Thorndike's first 1500,¹⁰ and the number of simple

⁶H. Y. McClusky, "A Quantitative Analysis of the Difficulty of Reading Materials," Journal of Educational Research, XXVIII (December, 1934), 276-82.

⁷W. S. Gray and B. E. Leary, What Makes a Book Readable . . . An Initial Study. Chicago: The University of Chicago Press, 1935.

⁸E. Dale, "A Comparison of Two Word Lists," Educational Research Bulletin, X (December, 1931), pp. 335-64.

⁹C. Washburne and M. V. Morphett, "Grade Placements of Children's Books," Elementary School Journal, XXXVIII (January, 1938) pp. 335-64.

¹⁰Thorndike, op. cit.

sentences in 75 sample sentences. Lorge's formula¹¹ and its later revision examined average sentence length in words, the number of prepositional phrases per 100 words, and the number of different "hard" words (i.e. not on the Dale list) per 100 words.¹² The Lorge formula made great strides in the direction of economy of effort. However, the Flesch formula,¹³ which appeared shortly after, proved even more popular, particularly in its revised form.¹⁴ The first form was based on average sentence length in words, number of affixes appearing and the number of personal references. The second form measured readability as a function of the number of syllables per 100 words and the average number of words per sentence. A further "human interest index" was also provided for. A still very popular formula which followed, the Dale-Chall,¹⁵ used as factors the number of words not appearing on a provided list of 3000, and average sentence length in words. The return to vocabulary factors tends to increase the time required to apply this approach. A very simplified approach is Gunning's

¹¹I. Lorge, "Predicting Reading Difficulty of Selections for Children," Elementary English Review, XVI (October, 1939), pp. 229-33. I. Lorge, "The Lorgeard and Flesch Readability Formulae: A Correction," School and Society, LVII (February 21, 1948), pp. 141-42.

¹²Dale, op. cit.

¹³R. F. Flesch, Marks of Readable Style: A Study in Adult Education. New York: Bureau of Publications, Teacher's College, Columbia University, 1943.

¹⁴R. F. Flesch, "A New Readability Yardstick," Journal of Applied Psychology, XXII (June, 1948), pp. 221-33.

¹⁵E. Dale and J. S. Chall, "A Formula for Predicting Readability," Educational Research Bulletin, XXVII (January 21, 1948), pp. 11-20.

Fog Index,¹⁶ which uses average sentence length and the percentage of words of 3 or more syllables. The Spache formula,¹⁷ designed for primary grade materials, uses techniques from several other studies, employing average sentence length and the number of words outside of Dale's list of 769.¹⁸ Several other studies look at such aspects as "abstractions," and shall not be discussed here.

The above review of representative tests of "reading difficulty" and comprehension represent practical solutions to practical problems. A test of readability should be objective, as well as simple and quick to apply. Several formulas looked at the presence of prepositional phrases, or the frequency of occurrence of simple sentences. However, no real attempt was made to provide for contextual influences, nor was any attempt made to perform detailed analysis according to descriptive grammars. Such an analysis would be cumbersome in the extreme, and would provide little useful information, as no "scale of relative difficulty" for such structures would exist. If tests of the difficulty of the various syntactic structures of language are to be devised, it must be possible to quantify the accuracy of the readers' responses to such material before any progress can be made.

¹⁶R. Gunning, The Technique of Clear Writing. New York: McGowan-Hill, 1952.

¹⁷G. Spache, "A New Readability Formula for Primary-Grade Reading Materials," Elementary School Journal, LII (March, 1953), pp. 410-13.

¹⁸Dale, op. cit.

The following section suggests a procedure which holds promise of accomplishing this task.

The Cloze Technique

Various measures have been used by experimenters to determine the relative facility of subjects' responses to structures of language. The simplest of these measures has been the subject's ability to reproduce what he has been exposed to. Another approach has measured time lags involved before accurate responses have been made. A third has involved the task of reproducing as many words as possible from a random list inserted after or between test structures. The first and last of these approaches is highly dependent upon what can be termed "short-span memory," and if an aural presentation is used, further perceptual problems arise. The time lag approach cannot differentiate clearly between the time required for perception and actual "decoding and encoding" time. Clearly, an approach which avoids dependence on short-term recall, and which reduces perceptual problems to a minimum is desirable. Further, if clear cut criteria for accepting or rejecting a response can be established, and if these criteria can be related to hypotheses concerning the nature of the structures of language, then many of the problems of much of the research reviewed in the previous section can be reduced if not eliminated.

One such approach which seems to hold promise has

been developed by Taylor,¹⁹ who has called it "Cloze procedure." Originally designed as a tool to gauge readability, this technique is simple in design, although somewhat time-consuming to score. The technique consists of deleting either every n^{th} word, or words of particular type from a passage, and requiring subjects to fill in the missing words. Missing words are replaced by blanks of equal length. Scoring can be on the basis of accepting only the exact word deleted, or accepting synonyms as well.

A number of studies have used the "Cloze" technique: for example, Coleman and Blumenfeld,²⁰ as well as Fagan²¹ have both utilized this approach, which seems to fulfill the requirements for an effective approach outlined above. Further, their application of this technique has been for the same purpose as the present study: to determine the relative difficulty of various structures of language, a function for which the technique was not originally developed.

As stated above, the "Cloze procedure" was meant to be a test of readability of materials. For example, the normal cloze procedure involves replacing every fifth word

¹⁹Wilson L. Taylor, "Cloze Procedure: A New Tool for Measuring Readability," Journalism Quarterly, XXX (Fall, 1953).

²⁰E. B. Coleman and J. P. Blumenfeld, "Cloze Scores of Nominalization and their Grammatical Transformations Using Active Verbs," Psychological Reports, XIII (1963), pp. 651-654.

²¹William F. Fagan, "An Investigation Into the Relationship Between Reading Difficulty and the Number and Types of Sentence Transformations." (Unpublished Doctoral dissertation, University of Alberta, 1969).

with a blank. Thus five alternative forms of a comprehensive "Cloze" test should be prepared, the first deleting the first, sixth, and eleventh words and so on, the second deleting the second, seventh, twelfth, and so on. Scores on the five forms could then be averaged (or totalled). A similar procedure applied to a second passage using the same subjects would then give an index of comparative readability for the two passages.

Unlike previous approaches to readability measurement, cloze procedure does not involve counting of elements nor determining the length of sentences. It rather seems to call upon the subject's sensitivity to contextual and structural clues surrounding the blanks. Writes Taylor, "It seems to measure whatever effect elements actually have on readability. And it does so at the same time it is also taking account of the influence of many other factors readability formulas ignore."²² In a later study, Taylor²³ began to examine the further applicability of "Cloze" and writes, "if the statement that a passage is 'readable' means that it is 'understandable,' then the scores that measure readability should measure comprehension too."²⁴ Taylor began to see further possibilities for "Cloze" as a measure of the reading--and

²²Taylor, op. cit., p. 417.

²³Wilson L. Taylor, "Recent Developments in the Use of 'Cloze Procedures'," Journalism Quarterly, XXXIII (1956).

²⁴Taylor, *ibid.*, p. 44.

possibly linguistic--abilities of individuals.²⁵ He states that success on a "Cloze" test seems to rest heavily upon the subject's grasp of passage meaning, an accomplishment which Taylor says involves ". . . general language facility, specific knowledge and vocabulary relevant to the material at hand, native ability to learn, attention, motivation and so on."²⁶ In order to examine this rather sweeping hypothesis, Taylor conducted a study in which he prepared a number of alternative "Cloze" forms on a single passage, forms from which words were deleted in three ways: an "any" test, deleting every n^{th} word; an "easy" test, deleting words found to be easy on previous tests; and a "hard" form, deleting hard or difficult words. These tests were administered to subjects along with standardized multiple choice tests of comprehension.

Paired distributions of scores on comprehension tests and on the three "Cloze" test forms were found to correlate highly significantly and positively. Further, test-retest reliability coefficients obtained on the three forms of the "Cloze" were as follows: "any" $-.88$; "easy" $-.80$; "hard" $-.84$. A further correlation performed as a check showed highly positive and significant results in comparisons of "Cloze" test scores and those on the Armed Forces Qualification Test.

²⁵Wilson L. Taylor, "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Attitude," Journal of Applied Psychology, XVI (February, 1957), pp. 19-26.

²⁶Taylor, 1957, op. cit., p. 19.

Strangely, not many researchers turned immediately to "Cloze" procedure as a measure of comprehension. Jenkinson,²⁷ one of the few who did, listed the following reasons for using it:

1. A logical analysis of the processes which are brought into play in the completion of the "cloze" test suggests both selected general and specific factors which would be involved. From the general point of view, it was reasoned that many of the types of thinking involved in reading as suggested by reading specialists, such as evaluating, judging, imagining, reasoning, and problem solving, might be employed in the completion of this type of test. Specifically, context would need to be used, and the meaning of individual words have to be fused together before the word required for the completion of the idea could be inserted.
2. It appears also that an individual's ability to complete a "cloze" test depends upon the extent to which he understands the meaning of the passage, and then on the various factors generally included in comprehension, as well as on the knowledge of vocabulary and general language ability.
3. Most reading tests require subjects to answer questions designed to measure the amount of knowledge obtained from passage. Such a method depends upon the ability of the subject to understand the question as well as the passage. Furthermore, it seems probable that the understanding which is gained by an individual in the ongoing reading process may differ from the comprehension which ensues when the reader is guided by questions. This criticism would appear to be even more apposite when multiple-choice questions are used.
4. Russell suggests that the process of concept formation "involves inductive thinking, and at least some deductive and creative thinking to clarify and verify the structure of the concept." Analysis of the "cloze" procedure reveals that these types of thinking are brought into play on the verbal level.
5. The "language" factor as it is designated by Burt which relates to words in context, and Thurston's word fluency factor appear to involve vital language abilities which are rarely measured in tests of reading

²⁷Marion D. Jenkinson, "Selected Processes and Difficulties of Reading Comprehension" (Unpublished Doctoral dissertation, the University of Chicago, 1957) p. 59.

comprehension. Both the ability to understand words in context and verbal fluency will obviously be tested by the "cloze" procedure.

The above quotation pinpoints several factors which the "Cloze" test examines which are vital to the present study. Among these are ability to imagine and reason; general language ability; word fluency; and ability to make use of contextual clues. Most critics would identify the first three of these to be clearly the mark of better or more mature writers, while the latter factor relates directly to awareness of the structures of language. Further discussion of these points will follow in the section on "The Instrument."

Rankin,²⁹ examining the rationale of "Cloze," quotes Taylor as saying:

A cloze test may be considered as a sample of message redundancy, because it samples the reader's ability to predict what word comes next at randomly chosen points in the article.³⁰

Rankin states that "Cloze" tests of "any word" form seem to correlate highest with I.Q.'s.³¹ He further describes a technique used by Jenkinson³² in which she interviewed students while they performed "Cloze" tests. She asked them to verbalize their thoughts while doing such a test, and found

²⁹Earl F. Rankin, Jr., "The Cloze Procedure--Its Validity and Utility," Eighth Yearbook of the National Reading Conference, 1959, pp. 131-144.

³⁰Rankin, *ibid.*, p. 135.

³¹Rankin, *ibid.*, p. 135.

³²Jenkinson, *op. cit.*

that students scoring high on the "Cloze" test,

. . . demonstrated significantly greater superiority in such characteristics as recognizing syntactical clues, sensitivity to style, fusion of separate meanings into ideas, recognition of implied meanings, verbal flexibility, knowledge of word meanings and language structure.³³

Fletcher³⁴ found the "Cloze" procedure useful in assessing readers' ability to use context, one of the primary skills such a test requires.

Ruddell³⁵ used the "Cloze" procedure as a measure of comprehension of passages using high and low frequencies of structures found common in the speech of grade 4 students by Strickland.³⁶ He found the "Cloze" to have a reliability of over .90, and achieved significant results indicating that passages high in the patterns of oral language were easier to read than passages low in such structures. Bormuth³⁷ pointed out the need for approaches to measuring readability which took cognizance of ability to read, not only individual words, but also sentences (Bormuth, p. 85). He further noted that such approaches should be applicable to everything

³³Rankin, *ibid.*, p. 141.

³⁴J. E. Fletcher, "A Study of the Relationships Between Ability to Use Context as an Aid in Reading and Other Verbal Abilities," (Unpublished Doctoral dissertation, University of Washington, 1959).

³⁵Robert B. Ruddell, "The Effect of Oral and Written Patterns of Language Structure on Reading Comprehension," The Reading Teacher, 18 (1965), pp. 270-275.

³⁶Ruth G. Strickland, "The Language of Elementary School Children: Its Relationship to the Language of Reading Textbooks and the Quality of Reading of Selected Children," Bulletin of the School of Education, XXXVIII, No. 4, Bloomington: Indiana University, July, 1962.

³⁷John R. Bormuth, "Readability: A New Approach," Reading Research Quarterly, I, (Spring, 1966), pp. 79-132.

contained in a book including indices, picture captions, and so on.

In a later study, Bormuth³⁸ used the full form of the "Cloze" test--that is, deleting every fifth word, and preparing five forms so that every word is eventually deleted--and concluded that the "Cloze" procedure was a useful measure of comprehensibility at the levels of words, independent clauses, and sentences.

Miller and Coleman³⁹ used various forms of the "Cloze" procedure, from only one word deleted, to forms where all words were deleted, and concluded, "cloze scores give a measure of readability over the whole range of difficulty" ⁴⁰

Mosburg, Potter and Cornell⁴¹ reported correlations ranging upward of .70 between "Cloze" test scores and scores on multiple-choice tests of comprehension. Skyler⁴² conducted similar research with similar results.

³⁸John R. Bormuth, "Cloze Readability Procedure," CSEIP Occasional Report Number I, (U. of California, Los Angeles, October, 1967).

³⁹G. R. Miller and E. B. Coleman, "A Set of 36 Prose Passages Calibrated for Complexity," Journal of Verbal Learning and Verbal Behavior, XI, (1967), pp. 851-854.

⁴⁰Miller and Coleman, *ibid.*, p. 854.

⁴¹L. Mosberg, J. C. Potter and R. K. Cornell, "The Relation Between Cloze and Multiple Choice Test Scores as a Function of Relative Paragraph Difficulty and Grade Level," Southwest Regional Laboratory Technical Report No. 6, 1968.

⁴²G. Skyler, "Comparison of the Cloze Procedure and the Short Answer Achievement Test in Determining Subject Matter Comprehension," Graduate Research in Education, V, (Fall, 1969), pp. 29-45.

One different use of the "cloze" procedure was demonstrated by Aquino, Mosberg and Sharron,⁴³ who used it and the Stanford Reading Achievement Test to gauge the readability of material in several fields: The correlations obtained were as follows:

| | | |
|-----------------|-------------------|------|
| Subject Area: | Science | .515 |
| | Human Interest | .612 |
| | TV-Movies-Theatre | .644 |
| Ability Groups: | High | .508 |
| | Medium | .651 |
| | Low | .610 |

In investigating the effect of deleting certain types of words, Bradley⁴⁴ extended research done earlier by Bickley, Weaver and Ford.⁴⁵ While they found that only the deletion of nouns had any great effect on comprehension, Bradley found that the deletion of either nouns or adjectives tended to lower scores on multiple-choice tests requiring exact answers. Bradley also found that the deletion of "function words"⁴⁶ did not lower scores significantly over full passages. Such results would seem to hold promise for "Cloze" tests minimizing semantic content.

⁴³Milagros Aquino, Ludwig Mosberg, and Marge Sharron, "Reading Comprehension Difficulty as a Function of Content Area and Linguistic Complexity," The Journal of Experimental Education, XXXVII No. 4 (Summer, 1969), pp. 1-4.

⁴⁴Michael Bradley, "Effects on Reading Tests of Deletions of Selected Grammatical Categories," Reading: Process and Pedagogy, 19th Yearbook of the National Reading Conference, Vol. 1, Process. pp. 87-93.

⁴⁵A. C. Bickley, W. W. Weaver, and F. G. Ford, "Information Removed From Multiple-Choice Item Responses by Selected Grammatical Categories," Psychological Reports, 23 (1968), pp. 613-614.

⁴⁶Charles C. Fries, The Structure of English, New York: Harcourt, Brace, and World, 1952.

Recent papers by both Bormuth⁴⁷ and Kirby⁴⁸ seem to hold out further promise for development of the "Cloze" procedure, and for quantifying its results.

In summarizing much of the research done using "Cloze" procedure up to the present time, Culhane⁴⁹ states,

One important fact that has emerged from these studies is that the cloze method is as good as, and in many ways better than, existing methods of teaching and testing comprehension.⁵⁰

Summary

Despite an awareness of the desirability of looking more closely at the structures of language,^{51,52} the earlier designers of formulas to gauge the difficulty of reading matter relied primarily upon matters of vocabulary and sentence length. The few formulas that did examine syntax did so in a very limited fashion. Further, no notice was taken of the effect of contextual clues upon facilitating reading.

⁴⁷J. R. Bormuth, "Empirical Determination of the Instructional Reading Level," International Reading Association Proceedings, XII (1969), Part 1, pp. 716-721.

⁴⁸C. L. Kirby, "Using the Cloze Procedure as a Testing Technique," International Reading Association Conference Proceedings, XIII (1970) Part 4, pp. 68-77.

⁴⁹J. W. Culhane, "Cloze Procedures and Comprehension," The Reading Teacher, XXIII (February, 1970), pp. 410-413.

⁵⁰Culhane, *ibid.*, p. 411.

⁵¹W. S. Gray and B. E. Leary, *op. cit.*

⁵²E. Dale and J. S. Chall, *op. cit.*

The "Cloze" test or procedure seems to bring many more factors of the reading act into play than previous tests of comprehension or reading formulas. Its reliance on the quality of redundancy in language examines, not only the subject's knowledge of lexical items but also his familiarity with syntactic structures, which are mutilated by the deletion process, and which must be restored.

"Cloze" tests seem to have the following advantages as tests of comprehension of passages of language:

1. they are easily constructed,
2. they are relatively easily scored, and
3. they seem to test many factors of language skill, including vocabulary, syntactic skill, and ability to use context clues.

The primary difficulty with "Cloze" procedure is that it is difficult to quantify its results, except by making comparisons with other tests of comprehension used on the same material.

For the purposes of this study, where results will not be generalized beyond the sample population, "Cloze" scores are sufficient to indicate both relative difficulties of the syntactic structures incorporated within the test, and the relative abilities of the subjects with the tasks which this test incorporates.

2. GRAMMARS AND LINGUISTICS

Grammars are descriptions of the substances and processes of languages: they are theories about how languages work. The least developed grammar would be a list of all the possible sentences in a language. Such a grammar would involve no generalization for rules whatsoever. In order to check the grammaticality of an expression, one would merely seek out a one-to-one correspondence with some item from such a list. Such a "grammar" (if it can, in truth, be called one) would be totally lacking in explanatory ability. At the other extreme would be a grammar which would capture every possible generalization of a language, and which would contain arrays of enormously complex rules and sub-rules. Such a grammar would be the ultimate in explanatory capacity with respect to its language or corpus, but would lose comprehensibility through its complexity.

The layman's concept of a grammar is as a set of rules for "accurate" usage; he wants to know whether a particular construction is "correct" (that is, whether it follows the idiom of what is considered to be "good" usage in the language community). He does not seek reasons nor rationales. He merely asks, "Do I talk good, or don't I?"

The linguist, on the other hand, seeks a maximum of explanatory power from any theory of language. He seeks to capture "significant generalization," to reduce ad hocness, and look for underlying principles. As this latter purpose is part of the rationale of this study, it seems necessary

for the writer to employ a modern grammar of relatively high explanatory capacity for this study. However, in order to justify the choice made, it is necessary to examine other and earlier grammars, and to assess their capabilities. Before turning to a survey of grammars, however, the necessity for using a grammar with explanatory capacity should be established.

Dominating the entire pedagogical problem of the teacher of English or language arts is the question of the nature of language itself. How can the language arts teacher set about to improve a skill which he himself only hazily understands, and which he can evaluate only in a highly subjective manner? Further, it seems as though some students are simply incapable of achieving what we might call higher levels of syntactic maturity or complexity.¹ Are some people inherently syntactically inept? This study deals with one facet of the psychological processes which occur during the events of language production and reception among human beings. Does the process of syntactic mapping and decoding of messages impose upon those messages, because of the actual syntactic shape which such messages take, greater or lesser degrees of complexity and/or developmental maturity?

Human language, amazingly flexible as it is, is also highly structured. Precisely how the human child acquires his ability with language is an unsolved problem, although a

¹John B. Carroll, Language and Thought, Englewood Cliffs, New Jersey, Prentice Hall, Inc., pp. 66-74.

number of suggestions have been made. For example, behavioristic paradigms of the language-learning process have been expounded and bitterly attacked.² It appears, on the surface at least, that some form of classical conditioning and mediating responses should account for some aspects of language acquisition.³ However, there are severe problems with the behavioristic approach. Not the least of these is the ability of persons to produce expressions which another speaker of that language will recognize as acceptable within the language in question even though the speaker (and even the "judge" of the expression) may never have encountered that particular expression before.⁴

The native speaker of a language acquires not only quite vast amounts of vocabulary over a short period of time, but also acquires (or possibly is inherently in possession of) a scheme for classification, organization, and concatenation of language, a mental grammar, a "mental reality underlying actual behavior."⁵ This capacity or ability, which all humans seem to share, processes linguistic components into expressions which are intelligible, which have meaning or truth value, and which can be communicated to another speaker

²Noam Chomsky, "A Review of B. F. Skinner's Verbal Behavior" (New York: Appleton-Century-Crofts, Inc., 1957) in Language, 35, No. 1 (1959), 26-58.

³Robert F. Terwilliger, Meaning and Mind, New York: Oxford University Press, 1968, pp. 46-91.

⁴Noam Chomsky, 1965, op. cit. pp. 47-62.

⁵Ibid., p. 4.

of that language, who somehow shares or has in common with the speaker sufficient of that same system (or mental grammar) to permit communication.

The exact physiological (or biological) nature of this so-called mental grammar is and must remain a mystery until scientists of the future do much more to unlock the secrets of the processes of the human brain. Consequently, it is almost impossible to make empirical statements about the processes of human language production except in terms of articulatory phonetics, electromyography, and the like. What we can do is to examine the output of the mental grammar, and attempt to characterize and classify such output. Such attempts are what we call "grammars." They deal, not with process, but with product.

One possible grammar of a language might be a simple list of all possible expressions of that language complete with all possible phonological patterns. Such an open-ended grammar obviously would be cumbersome to the point of infinity even for a simple language. However, some people have believed that the human brain was a tabula rasa, and that the pattern of language acquisition was a simple memorization of everything heard combined with a process of recall and reconstitution at the moment of desired production. One of the primary arguments against such a theory is the fact of recursion in natural languages. There is no expression in a language such as English which cannot be added to by means of modification, conjunction, and so forth. Thus, while at

any time the lexicon, syntactic rules, and phonological rules of a grammar might be said to be finite, the potential output of such a grammar, because of the feature of recursion, is truly open-ended. Consequently, a "list" grammar could never approach total adequacy. As a model of the mental grammar of the child, however, the list would seem to correspond to fact, especially for very young children. However, it seems as though the list must fail as a model as soon as the process of linguistic creativity of original production of expressions outside the child's experience begins.

Beyond the "list of expressions" lie grammars which separate the lexicon and the syntactic components. Dictionaries of English (which only give the semantic readings of words, and which do not specify occurrence restrictions beyond giving an example or two) have been available as early as 1604.⁶ But the area of greatest change in approach to grammars, particularly during the past fifteen years or so, is within the syntactic component.

The first real grammars of English were the so-called pedagogical or prescriptive grammars, which began to be popular in the early eighteenth century. Such grammars attempted to establish rules of usage, usually based upon grammars of Latin. They classified words into "parts of speech," the definitions of which were far from consistently

⁶James R. Gaskin and Jack Suberman, A Language Reader for Writers. Englewood Cliffs, New Jersey: Prentice-Hall, 1966, p. 34.

being functional.⁷ Such grammars accomplished a great deal toward creating a "standardized" language, but are of little interest to the investigator of the process of language production, or to the linguist.

Another approach was termed the "historical approach" of the neogrammarians. This method sought to catalogue and to classify, in as complete a manner as possible, all of the structures and items of the language in question, as well as the history of their derivations. This grammatical approach has its merit in its thoroughness; however, it results in grammars of considerable length and complexity.⁸ Further, such grammars often do not recognize or characterize certain of the generalities of the language in question. They are purely descriptive. While such grammars embody vast amounts of information about a language, they develop few ideas of interest to the investigator of the syntactic processes of language, except that they might serve as bodies of information--that is, as a corpus.

Another somewhat more difficult to classify school of linguistics is associated with the name of Leonard Bloomfield,⁹ probably the most influential figure in the entire field of linguistics from the 1930's up until about

⁷For example, Robert Lowth, A Short Introduction to English Grammars, (London, 1762), Reprinted: Monston, England: Scholar Press, 1962.

⁸For example, Otto Jespersen, A Modern English Grammar on Historical Principles, New York: Barnes and Nobles, 1956.

⁹Leonard Bloomfield, Language, New York: Holt, Rinehart and Winston, 1933.

1955. The "Bloomfieldian" linguists were first concerned with the sounds of language, although they addressed syntax in terms of "immediate constituent analysis" as well. They examined such features as intonation and stress patterns, and their published works seemed far more concerned with phonology. Early works such as Nida's thesis of 1943¹⁰ did not receive much attention, and so the rigorous study of syntax did not take on any great impetus from the American descriptive grammarians.

During and after World War Two, students of language began to become more interested in the syntactic component. This impetus was at least in part the result of a greater interest in foreign language learning brought about by the war. It was realized that new approaches to syntax might provide insights to the workings of the structure of languages, insights which might at least prove to be a starting point for studies into the very bases of language learning. Thus was born the structuralist school of linguistics, tracing back to Bloomfield, but chiefly represented by Charles Fries.¹¹ While Fries was not the originator of the concept, his work focused the attention of grammarians on the patterns of sentences, and on more accurate characterization of word classes.

¹⁰Eugene Albert Nida, A Synopsis of English Syntax, Norman, Oklahoma, Summer Institute of Linguistics, 1960.

¹¹Charles C. Fries, The Structure of English, New York: Harcourt, Brace and World, 1952.

A type of grammar similar to the type upon which the present study is based was first suggested by Zellig Harris.¹² While Harris' early formulations were particularly limited with respect to the scope of the transformational process beyond simple or singulary operations, the ideas contained in the basic concept were of sufficient germinal value to stimulate further work within the transformational framework.

The turning point for this theory of grammar was Noam Chomsky's formalization and extension of his earlier work, Syntactic Structures.¹³ While Harris had talked of the feature of recursiveness, Chomsky specifically suggested that only a generative-transformational model of grammar which took cognizance of the recursive properties of natural languages could serve as a first approximation for the "competence" element of the language-producing ability of the native speaker. This is not to say that Chomsky was putting forward his model as having any correspondence to what went on in the mental processes of the speaker.¹⁴ Rather, Chomsky merely said that only a model of grammar which incorporated an element providing for recursion and the embedding of one sentence within another could approach observational adequacy; that is, could produce accurately the data of a corpus of a

¹²Zellig S. Harris, Methods in Structural Linguistics, Chicago: The University of Chicago Press, 1951.

¹³Noam Chomsky, Syntactic Structures, The Hague: Mouton and Co., 1965.

¹⁴Chomsky, *ibid.*, p. 48.

natural language such as English. Of course, whether Chomsky's model is the only model capable of doing this is extremely doubtful. That Chomsky's claim in this respect is too strong is certain. The generative-transformational element is probably only one of a number of possible alternative formulations of the syntactic component. However, Chomsky's has at least been made fairly explicit.

Of particular interest is Chomsky's discussion of two other possible models of language-producing "devices"--a Markov-process finite state automaton, and a phrase structure grammar.¹⁵ The "finite state" model of grammar assumes a unidirectional process of language production (in written English, an output stream moving from left to right), the device moving from one "state" to the next producing words as it goes, optionally repeating any one state, and so forth. At times, alternate paths are provided thus:¹⁶

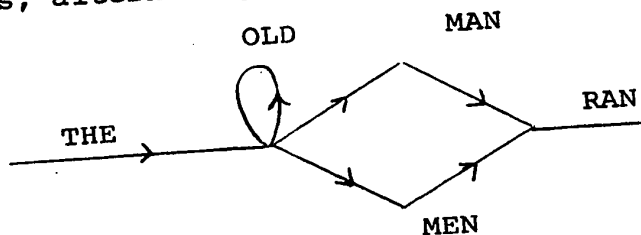


FIGURE I
FINITE STATE MODEL

The optional "loop" in this diagram provides for "the man ran," "the old man ran," "the old, old men ran," and so forth. At

¹⁵Noam Chomsky, Syntactic Structures, The Hague: Mouton and Co., 1957, p. 18-48.

¹⁶Noam Chomsky, op. cit., p. 19.

naive first glance, such a model seems reasonable: we do produce sentences as strings of words moving along in chronological (or left-to-right spatial) order. Chomsky's argument against such a grammar is that natural languages incorporate contingencies and dependencies within sentences; it is essential that some process for "looking back," or anticipation of what is to come be provided in the grammar of a natural language. Chomsky claims that the finite state automaton moving mechanically from one state to another, cannot incorporate such a process.

Chomsky's second suggestion was a "phrase structure" grammar, a model which provides a description of how a sentence is produced by means of constituent analysis. Symbols are used for classes of elements: for example, "S" stands for "sentence," "NP" for "noun phrase," "VP" for "verb phrase," "Det" for "determiner," "N" for "noun," "V" for "verb," and so forth. A phrase structure grammar takes the form of a series of "rewrite rules," which specify that a particular symbol is to be rewritten as another symbol or series of symbols. An arrow (\rightarrow) means "is to be rewritten as . . ." Only one symbol may appear on the left of an arrow; that is, only one element may be rewritten at a time, although it may become one or more other elements. Lexical insertion may take the form of rewrite rules, or a simple lexicon may be provided from which substitution for symbols may be made. A typical sentence might be produced as follows:

1. S -----> NP + VP
2. VP -----> V + NP
3. NP -----> Det + N
4. Det -----> the
5. N -----> dog, horse, man, . . .
6. V -----> bit, struck, loved, . . .

The above rules would give the following as one possible complete derivation:

S
 NP + VP
 NP + V + NP
 Det + N + V + NP
 Det + N + V + Det + N
 the + man + V + Det + N
 the + man + loved + Det + N
 the + man + loved + the + N
 the + man + loved + the + dog

The following phrase marker or "tree diagram" represents the above derivation:

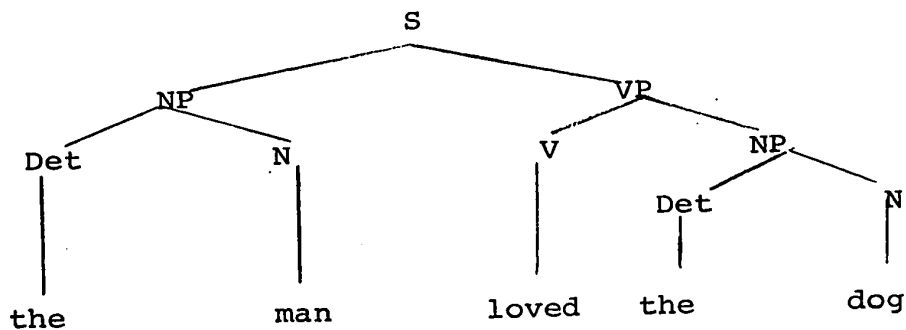


FIGURE II
 PHRASE MARKER

The final string of symbols before lexical insertion is called a "terminal string," because none of the symbols within it can be rewritten further. Notice that different selections of lexical items might have yielded "the horse loved the dog," "the man bit the horse," or "the dog struck the man." Notice further that tree markers do not give any indication of the order of derivations, as complete written derivations do. The manner in which the various symbols are rewritten is somewhat similar to "immediate constituent analysis," although not identical because of the requirement in "I.C." analysis that each division (or "cut") be just that--one element must become two, and only two. Another thing to notice is that a phrase structure grammar makes certain generalizations about sentence elements (they can be rewritten as other things) and about entire sentences (several sentences may have identical terminal strings underlying them). The order of concatenation of elements also captures the concept of word order.

Chomsky does not deny that it would be possible to write a phrase structure grammar which would generate all possible sentences. His argument is that such a grammar would be excessive in size and clumsiness. As examples of situations which phrase structure grammars cannot handle well, he cites the process of conjunction, the presence of discontinuous elements, and the special requirements of the process of passivization. Given two sentences, "Tom went to Chicago" and "Bill went to Chicago," we would like to be able

to say, "Tom and Bill went to Chicago." Phrase structure grammar is not capable of relating these two shorter sentences and deleting the common elements. In changing the sentences, "Sam gave John the book" to "Sam has given John the book," one element, "have + en," has been added, the affix "en" having the effect of changing "gave" to "given." However, no phrase structure grammar can capture the generality of affix movement, a process which greatly simplifies the analysis of the verb phrase in English. As a third example of "phrase structure" problems, Chomsky cites the many restrictions which passive forms impose: verbs must be transitive. Further, the acceptability of "John plays golf," and the unacceptability of "Golf plays John" requires a set of rules to exclude the latter. But if the auxiliary includes be + en, all such restrictions fail; they need to be restated in reverse order. While such a duplication could be included in a phrase structure grammar, it is just further evidence of the clumsiness of such models. States Chomsky,

This inelegant duplication, as well as the special restrictions involving the element be + en, can be avoided only if we deliberately exclude passives from the grammar of phrase structure, and reintroduce them by a rule such as:

(34) If S_1 is a grammatical sentence of the form

$$NP_1 - Aux - V - NP_2,$$

then the corresponding string of the form

$$NP_2 - Aux - + be + en - V - by + NP_1$$

is also a grammatical sentence.

For example, if John - C - admire - sincerity is a sentence, then sincerity - C + be + en - admire - by

+ John (which by (29) and (19) becomes "sincerity is admired by John") is also a sentence.¹⁷

Thus does Chomsky introduce the notion of the model of grammar which he favours, a "transformational" model, which he describes as having three parts: a phrase structure base, a transformational structure, and a morphophonemic component (which gives sentences their final form).¹⁸ In speaking of transformations, Chomsky notes that "we must define an order of application" for them.¹⁹ and we must distinguish between optional and obligatory transformations.²⁰ He further defines "kernel" sentences as, "the set of sentences that are produced when we apply obligatory transformations to the terminal strings of the (phrase structure) grammar."²¹

In making specific his position at the time of writing Syntactic Structures²² with respect to the psychological reality of transformational grammars, Chomsky says:

A grammar does not tell us how to synthesize a specific utterance; it does not tell us how to analyze a particular given utterance. In fact, these two tasks which the speaker and the hearer must perform are essentially the same, and are both outside the scope of grammars of the form (35). (Chomsky's tripartite transformational model). Each such grammar is simply a description of a certain set of utterances, namely, those which it generates.²³

¹⁷Chomsky, op. cit., p. 43.

¹⁸Chomsky, op. cit., p. 46.

¹⁹Chomsky, op. cit., p. 44.

²⁰Chomsky, op. cit., p. 45.

²¹Chomsky, *ibid.*

²²Chomsky, op. cit.

²³Chomsky, op. cit., p. 48.

From the time of Syntactic Structures to the present, Chomsky has continued to develop his transformational theory.^{24, 25, 26, 27, 28, 29} His book Aspects of the Theory of Syntax³⁰ has been one of the most influential of his writings, and the theory contained therein, together with some modifications in the base component, provide the basic grammar of the present study. The "Aspects" model and suggested modifications to it will be discussed in detail in the chapter dealing with "The Grammar of the Study; however, a brief sketch will also be given in this review. While Chomsky has turned more strongly toward the study of phonology,³¹ his work in both syntax and phonology has followed the pursuit of what might be called "linguistic universals," processes and strategies which underly all human

²⁴Noam Chomsky, "Some Methodological Remarks in Generative Grammar," Word 17, pp. 219-239.

²⁵Noam Chomsky, "A Transformational Approach to Syntax," in A. A. Hill (ed.), Proceedings of the 1958 Conference on Problems of Linguistic Analysis in English, pp. 124-148, Austin, Texas, 1962.

²⁶Noam Chomsky, Current Issues in Linguistic Theory, The Hague: Mouton, 1964.

²⁷Noam Chomsky, Aspects of the Theory of Syntax, Cambridge, Mass.: The M.I.T. Press, 1965.

²⁸Noam Chomsky, Cartesian Linguistics, New York: Harper and Row, 1966.

²⁹Noam Chomsky, "Topics in the Theory of Generative Grammar," in T. Sebeok, ed., Current Trends in Linguistics, 3: Linguistic Theory, Bloomington: Indiana University Press, 1966.

³⁰Chomsky, op. cit.

³¹For example, Noam Chomsky and Morris Halle, The Sound Pattern of English, New York: Harper and Row, 1968.

language. Indeed, transformational theory has been successfully applied to the analysis of many other languages.^{32,33,34}

With respect to applications of transformational principles to English syntax, it would be difficult to review the literature exhaustively. The following survey lists representative texts and applications up to the present.

Lees³⁵ presents a very thorough approach to the analysis of the "noun phrase." This work is the first fairly extensive attempt at formalizing what might be called a workable transformational grammar of English.

Miller and Chomsky³⁶ examined various models of language users, and noted particularly the need for complete systems of "features" on lexical items. This analysis was reflected in Chomsky's Aspects of the Theory of Syntax,³⁷ published shortly thereafter. At about the same time,

³²Gary Dean Prideaux, The Syntax of Japanese Honorifics, The Hague: Mouton, 1970.

³³F. B. Deluze-Mommeja, "An algorithm d'analyse de la structure des phrases," Etudes de Linguistique Applique, 4(1966), 79-93.

³⁴Milka Ivie, "Non-admissible Determiners in Slavic Languages," Proceedings of the Ninth International Congress of Linguists, The Hague, 1966, pp. 476-479.

³⁵Robert B. Lees, A Grammar of English Nominalizations, Bloomington, Indiana: Research Centre in Anthropology, Folklore, and Linguistics, 1960.

³⁶G. A. Miller and N. Chomsky, "Finitary Models of Language Users," in R. D. Luce, R. Bush, and E. Galanter (eds.), Handbook of Mathematical Psychology, Vol. II, Ch. 13, pp. 419-492. New York: Wiley, 1963.

³⁷Chomsky, op. cit.

Katz and Postal³⁸ put forward a proposal for integrating grammatical and semantic descriptions of language. Chomsky's early model of a transformational grammar made no real provision for the many semantic problems of syntax. For example, an adequate grammar must be able to distinguish between the acceptable, "John plays golf," and the ungrammatical, "Courage admires Susie." The main problem of Katz and Postal is where their model of the semantic component of a grammar should apply to the Chomsky generative-transformational model of language which they adopt. They state that if the various transformations which might be applied to a phrase marker have the effect of changing meaning, the combined semantic-syntactic model will be necessarily very complex, although it could be made descriptively adequate. Katz and Postal propose that, by means of incorporating dummy symbols such as Q (question), IMP (imperative), and NEG (negative) into the base component of the model which they adopt, they would be able to introduce their semantic component at one level (namely, at the "deep structure" level) and consequently greatly simplify their combined model.

The syntactic model to which Katz and Postal were addressing themselves was not, in fact, the earlier Chomsky model, but that which Chomsky (also working at M.I.T.) put forward in Aspects of the Theory of Syntax. The Aspects

³⁸J. J. Katz and P. Postal, An Integrated Theory of Linguistic Descriptions. Cambridge, Massachusetts: The M.I.T. Press, 1964.

model proposes a process illustrated in Figure III. This model consists of a base component, a semantic component, a lexicon, transformational rules, and phonological rules. The base consists of:

1. branching or phrase structure rules, and
2. subcategorization rules, which are further divided into
 - a. context-free rules, and
 - b. context-sensitive rules.

The context-sensitive rules are of two types:

1. strict subclassification rules, and
2. selectional rules.

The derivation of a sentence through this base component results in a pre-terminal string of symbols, each having attached matrices reflecting both inherent and contextual features. Then, with reference to "readings" from the semantic component, items from the lexicon are inserted in place of the symbols with their attached feature matrices, the insertions being governed by those matrices and the context of the pre-terminal string itself. The resulting "terminal string" then passes through the transformational component, where it becomes the "surface structure" of the sentence. The phonological component converts surface structure into the sounds of spoken language.

Note that the Aspects model, in accordance with the theories of Katz and Postal,³⁹ postulates that the "deep structure" of the sentence--the product of lexical insertion

³⁹Katz and Postal, op. cit.

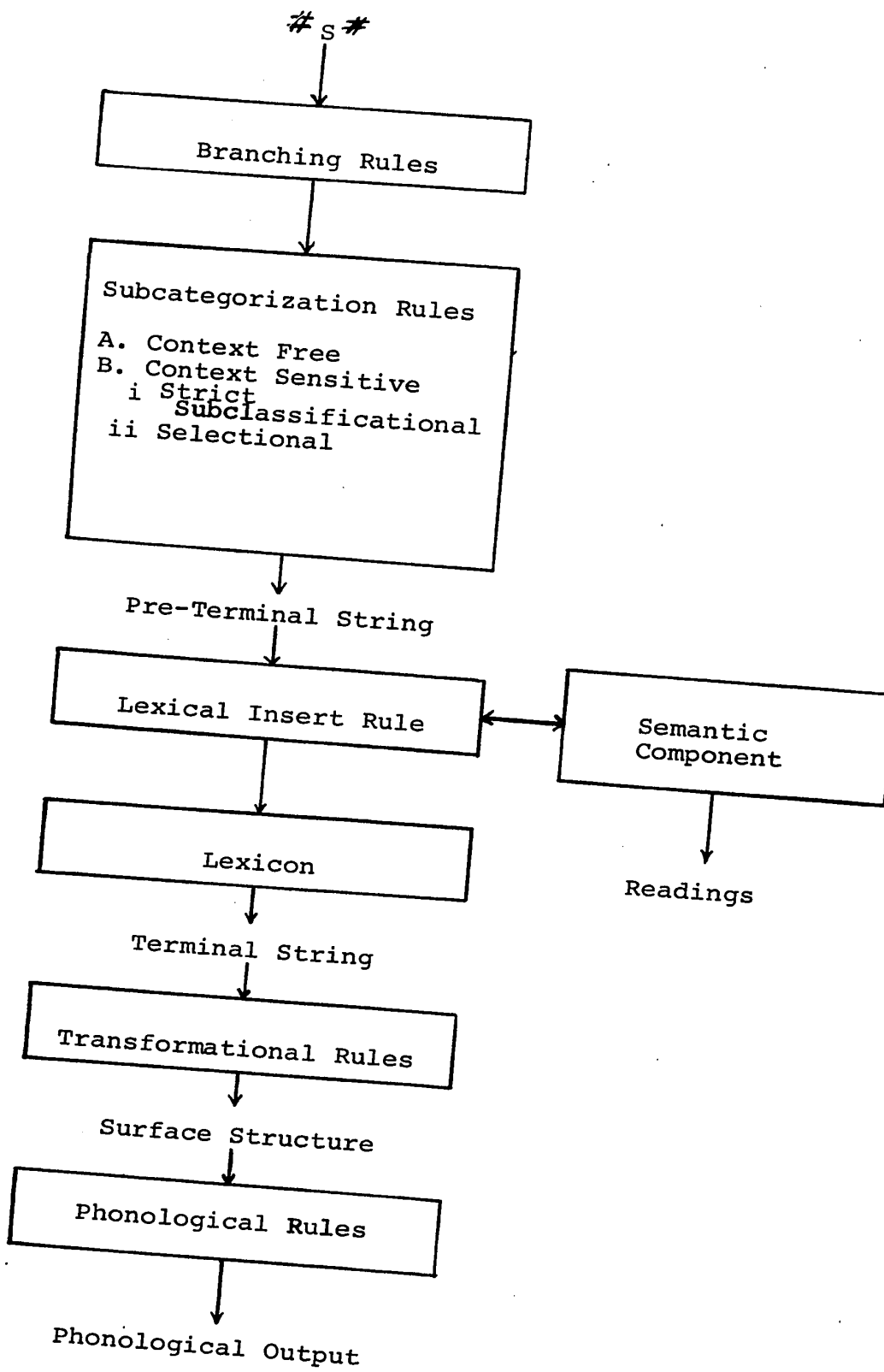


FIGURE III
CHOMSKY'S ASPECTS GRAMMAR

prior to transformational processes--carries the entire semantic load. That is, the process of application of optional transformational rules has no semantic effect. The accuracy of this claim has since been disputed, and will be further discussed.

The derivation of sentences using the Aspects model makes one fact abundantly clear: a great amount of redundancy appears within feature matrices. Take, for instance, the matter of subject-verb agreement. It would be possible to specify precisely, within the matrix of verb features, the type of noun which could be used as subject (and as object); or the nouns could carry specifications as to what kind of verb could appear with them. While the choice is arbitrary, it is somewhat more economical to place environmental restrictions on the verb element. Thus the feature matrix under the verb element in the sentence, "The man admires honesty" would be something like this:

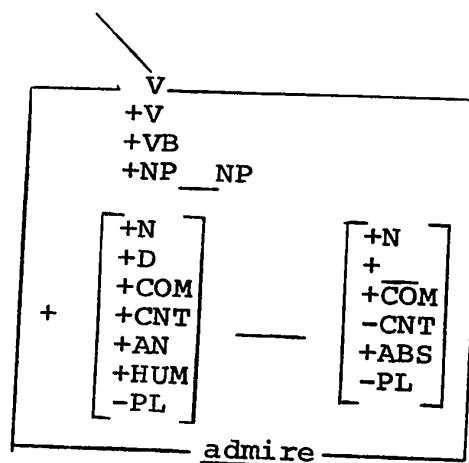


FIGURE IV

A FEATURE MATRIX

In the diagram, COM stands for "common," CNT stands for "count," AN for "animate," HUM for "human," PL for "plural," and ABS for "abstract." While such complex feature matrices are functional, they are awkward, and excessively redundant. They do represent, however, a means of expressing the existing dependencies and contingencies which exist semantically within natural languages such as English. Thus, if it is possible to have a feature such as +PL in a grammar as an inherent feature, it is also possible to list only singular forms in the lexicon with the feature [+PL]. A low order rule would have to change "dog" + PL into "dogs," "horse" + PL into "horses," and "man" + PL into "men." Some complications arise here, but the other alternative is to list both singular and plural forms in the lexicon. Indeed, a great expansion of the lexicon to cover such situations is precisely what has been suggested.⁴⁰ Transformational theorists, however, prefer to approach the notion of plurality in the form of a single simple rule or battery of such rules, thereby capturing a "linguistically significant generality."

It is possible to reduce the repetitiveness of many features in the "Aspects" model by means of a set of lexical redundancy rules, which stipulate, for example, that a noun marked (+HUM) would also be understood to be (+AN). Further, the proliferation of symbols within the verb feature matrix can be greatly reduced by moving contextual restraints into

⁴⁰James D. MacCawley, "The Role of Semantics in a Grammar," in Emman Bach and Robert T. Harms (eds.), Universals in Linguistic Theory. New York: Holt Rinehart & Winston, Inc., 1968, pp. 125-170.

the lexicon. Such a "post-Aspects model," (the model selected for this study), is considerably less redundant than the earlier model, although it retains its basic structure. Further discussion of the rationale for the selection of this "post Aspects model" is presented in Chapter III.

The Chomsky transformational model of grammar has not been without critics.^{41,42} The investigator is prepared to let research evidence either support or disprove any such theories. It is hoped that the present study will provide some useful evidence.

Even among modern linguists who favor some form of transformational theory, however, there have been objections to Chomsky's model, and proposals for alternatives. One such alternative has been called "case grammar." Case grammarians such as Fillmore⁴³ believe that the concept of "case" should be a part of the base component of grammars of every language.⁴⁴ Fillmore is attempting to solve the problem of combining syntactic and semantic relationships, the problem which Chomsky approaches via his "features analysis." Writes Fillmore,

I am going to suggest below that there are many semantically relevant syntactic relationships involving nouns and the structures that contain them, that these relationships . . . are in large part covert, but are nevertheless empirically discoverable, that they form a specific set, and that observations made about them will turn out to have considerable cross-linguistic validity.⁴⁵

⁴¹Carl A. Lefevre, Linguistics, English, and the Language Arts. Boston: Allyn and Bacon, Inc., 1970, pp. 357-361.

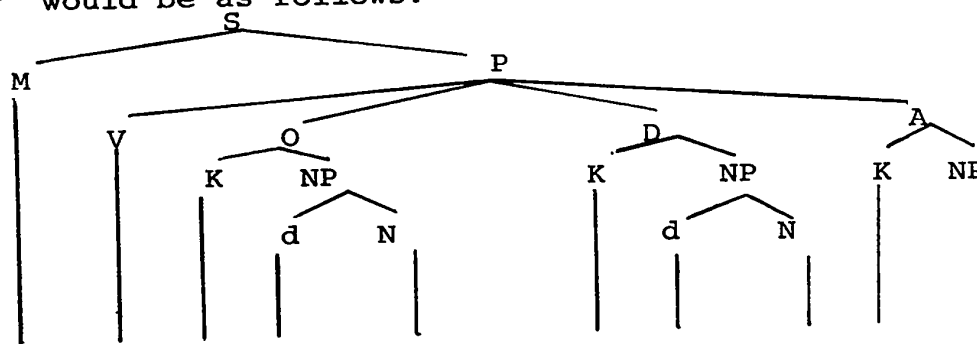
⁴²George Lakoff, "On Generative Semantics" (Unpublished manuscript).

⁴³Charles J. Fillmore, "The Case for Case," in Emman Bach and Robert T. Harms (eds.), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston, Inc., 1968, pp 1-88.

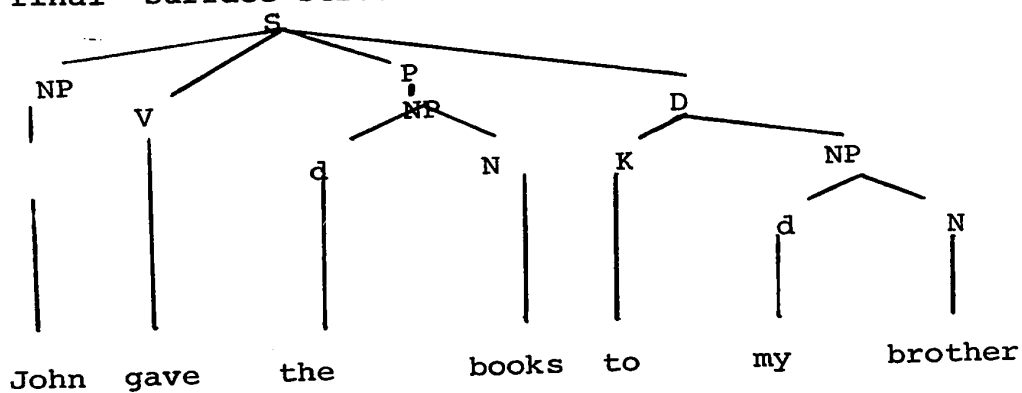
⁴⁴Fillmore, *ibid.*, p. 2.

⁴⁵Fillmore, *ibid.*, p. 5.

According to Fillmore, "The sentence in its basic structure consists of a verb and one or more noun phrases, each associated with the verb in a particular case relationship."⁴⁶ Case grammar theory sees each noun as having some kind of preposition (represented symbolically by "k" in the diagrams to follow) associated with it. Thus the "deep structure" case grammar phrase marker for "John gave the books to my brother" would be as follows:⁴⁷



Past give \emptyset the books to my brother by John
 The final "surface structure" form would be as follows:⁴⁸



⁴⁶Fillmore, *ibid.*, p. 21.

⁴⁷Fillmore, *ibid.*, p. 35.

⁴⁸Fillmore, *ibid.*, p. 36.

In particular, the assignment of structure to prepositions and their associated structures is difficult in "normal" transformational theory, but is handled very effectively by case grammars. As can be seen, however, it is a long way from case grammar deep structures to the intuitive level of language. One very large study, however, has recently opted for the use of a case grammar model.⁴⁹

Another approach typified by the work of McCawley⁵⁰ and Lakoff⁵¹ has been called "generative semantics." One postulate of this theory is that sentences having the same truth value should be derived from the same underlying structures. (Either that, or the entire concept of an underlying "deep structure which carries the complete semantic load of the sentence" must be revised.) For example, a generative semanticist would claim that "kill" and "cause to become dead" must have a common source. Supposedly, "become dead" would underlie "die," and then "cause to die" would be interpreted as "kill." Lakoff⁵² argues that there is no basis whatsoever for Chomsky's level of "deep structure," and there

⁴⁹(UCLA English Syntax Project "USEP"--p. 445.--Opt/Case Grammar.)

⁵⁰James D. McCawley, op. cit.

⁵¹George Lakoff, "On Generative Semantics," in Danny D. Steinberg and Leon A. Jacobovits, (eds.), Semantics: An Interdisciplinary Leader in Philosophy, Linguistics, Anthropology and Psychology. London: Cambridge University Press, 1969.

⁵²Lakoff, *ibid.*, p. 51-53.

exists excellent evidence as to the coalescence of syntax and semantics throughout the process of derivation. Thus he claims that the burden of proof of the existence of deep structure lies with those who have proposed it, and have then proceeded to assume it as "given."

Lakoff's definition of the position of the "generative semanticists" is "that syntax and semantics cannot be separated and that the role of transformations, and of derivational constraints in general, is to relate semantic representation and surface structures."⁵³ While Lakoff and his associates have turned up examples to support the above theory, it must be admitted that their attempts at making their theories explicit have not approached adequacy as yet. Consequently, Chomsky's model is still held to be viable by many, although it has been made subject to much patching and shuffling about. We are loath to give up what seems neat and tidy for a return to the chaotic realities of natural languages.

There is no doubt that theories of transformational grammar are undergoing a very lively period of development, and a healthy recognition of problems is being evidenced. However, one wonders if the "rule writers" are sometimes going too far, particularly with respect to inferences. Indeed, the word "rule" suggests a regularity to natural languages when in fact the term is used by linguists themselves only in reference to their closed grammars. This distinction should be made clearer.

⁵³Lakoff, *ibid.*, p. 65a.

The transformational linguist, in looking at languages, has become aware of a repetitive--almost mathematical set of regularities in what he observes. The system of tags and nomenclatures developed by the transformational linguist is useful; they help to codify certain of these regularities of language, and to focus attention upon the interrelations which exist within units of speech. For example, the transformationalists' development of the auxiliary system of English and how it operates can be most interesting and revealing to a speaker of English, who has used that system unconsciously and accurately for years. Yet research reveals that deliberately complicating the auxiliary system of a sentence does not, contrary to what transformational theorists would predict, make a sentence more difficult to comprehend or process.⁵⁴ The entire system seems so ingrained, its use so automatic, that we use it as naturally as we breathe. Consequently, here we have an instance which seems to indicate no relationship between a very economical and elegant transformational description of a state of affairs within a grammar and the empirical world of language performance.

The implications of such findings will be further discussed in the portion of this chapter dealing with psycholinguistics and the reality of the transformational process, but in summarizing the criticism of transformational

⁵⁴G. A. Miller and K. McKean, "A Chronometric Study of Some Relations Between Sentences," Quarterly Journal of Experimental Psychology, XVI, pp. 297-308.

grammarians, the writer believes it would be accurate to say that the critics are most bitter when the transformationalist makes the leap from the empirical "what is" to the philosophical "what I would like it to be," and implies psychological reality while in mid-air.

Beyond transformational theories of grammar are several interesting approaches. Pike relates a theory of "tagmemics."⁵⁵ Originally published in three volumes in 1954, 1955, and 1960, this work proposes a unit of grammatical analysis defined in terms of both meaning and structure, the tagmeme, that relates to an arrangement, an order, or a string of linguistic items. A tagmeme can also be a class of grammatical forms that function in a particular way. The tagmemic approach uses a "slot and filler" description whereby "slots" may be "filled" by any member of a particular tagmemic class.

Pike believed that thought only becomes articulated by means of language,⁵⁶ and cited both Sapir⁵⁷ and Vygotsky⁵⁸ as holding similar views. Thus he believed that the syntax

⁵⁵Kenneth L. Pike, Language in Relation to a Unified Theory of the Structure of Human Behavior. New York: Humanities Press, 1964

⁵⁶Kenneth L. Pike, College Composition and Communication, May, 1964, p. 83.

⁵⁷Edward A. Sapir, Language: An Introduction to the Study of Speech. New York: Harcourt, Brace, 1921, p. 15.

⁵⁸L. S. Vygotsky, "Thought and Speech," Chapter VII of Language and Thought, tr. Helen Kogan, Eugenia Hanfmann and Jacob Kasanin, in Psycholinguistics: A Book of Readings, ed. Sol. Saporta. New York: Holt, Rinehart and Winston 1961, p. 514.

and rhetoric of a paragraph would imply an analysis of the patterns of thought of the writer.

Another approach, stratificational grammar, has been put forward by Gleason⁵⁹ and Lamb.⁶⁰ One of the purposes of this school of grammar is to attempt to deal with units of language beyond the sentence. Language is viewed in terms of communication theory, and is seen as having three major "levels" (semology-dealing with meaning, grammar, and phonology) plus interconnecting "strata." A complete stratificational grammar would make explicit all of the various elements within each strata, plus the interrelationships existing between them. The great complexity of the task confronting stratificational grammarians has undoubtedly slowed their progress toward a rigorous theory.

Summary

To summarize the foregoing section, grammars have undergone changes not only in form, but also in philosophy. In truth, the earliest forms of prescriptive grammars of English are still with us in the desire of parents and pedagogues to have children speak and write "correct" English. The great descriptive grammars of English are, for the most part, just as accurate and comprehensive today as when they were written. The sentence patterns of structural grammarians continue to be in use, and transformational grammarians continue

⁵⁹H. A. Gleason, Jr., "The Organization of Language; A Stratificational View," Report of the Fifteenth Annual (First International) Round Table Meeting on Linguistics and Language Studies, ed. C. I. J. M. Stuart, Washington: Georgetown University Press, 1964.

⁶⁰Sydney M. Lamb, Outline of Stratificational Grammar, Washington: Georgetown University Press, 1966.

their search for "significant generalities" that they hope might give clues as to underlying thought patterns. It is not surprising that grammars have evolved on divergent paths when one considers their varying ends and philosophies. All that they have in common is their corpus, the English language

The best approach, perhaps, is that of the eclectic, using whatever is useful in a particular situation, and remaining open to new developments. The present study, which postulates that certain of the structures of English are more complex (and thus less easy to use or understand) than others, required some systematic nomenclature for identifying a number of the syntactic structures common in English. The nomenclature and organization chosen has been that of a post-Aspects grammar of the type discussed by Chomsky. The grouping of structures into embedding, simple, conjoining, position shift, and deletion groups has been done so that results could be compared with previous research, particularly that of Fagan.⁶¹ The choice of items or structures based upon a transformational model also permits this study to add to investigation into the Derivational Theory of Complexity, and the psychological reality of the transformational process. One point with respect to the grouping employed must be made clear, however. A so-called "simple" transformation such as the passive involves considerably more in the way of separate operations than, say, the process of negation. A T-Passive (passive transformation) requires addition of two elements

⁶¹Fagan, op. cit.

(a form of the verb "to be" plus "by") and the exchange of positions of the NP's. Contrast this with the addition of a pre-nominal adjective, supposedly an embedding transformation involving processes of relativization of a constituent sentence, deletion of elements, and finally a front-shifting of the remaining adjective. At the level of surface structure, this very complex embedding process takes the form of simply "plugging in" a single word. One wonders if linguists, in trying to get more mileage out of a transformational process required elsewhere (the relative transformation), haven't gone too far. Might not a recursive (ADJ) element in a prenominal position in the base be closer to reality? The answer to this question, which seems a logical one, is that transformational grammarians are not concerned with "performance levels" of language and their relative simplicity or complexity. They are working within their own formal systems, their purpose being to avoid ad hoc-ness and to seek as many generalities and commonalities as possible.

Some of these problems in transformational theory will be further discussed in the chapter dealing with the grammar of the study, but they also receive some attention in the section to follow dealing with psycholinguistics and the reality of the transformational model.

3. PSYCHOLINGUISTICS AND TRANSFORMATIONAL THEORY

Introduction

One of the persistent problems which have clouded the psycholinguistic horizons for the past ten years or so is the confusion surrounding the psychological reality of generative transformational processes. Again and again various theorists remark that their grammars are closed, formal systems designed to mechanically produce grammatical sentences. For example, Chomsky, who has made the above point often, writes,

. . . No doubt, a reasonable model of language will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer's knowledge of the language; but this generative grammar does not, in itself, prescribe the character or functioning of a perceptual model of speech production.¹

The problem lies in the fact that the only "generative grammar" that most people are familiar with is Chomsky's model, or something similar. Chomsky's generative-transformational model is put forward as capable of being a "competence" model of English grammar. Thus, questions arise when he writes on the page following the above quotation:

There seems to be little reason to question the traditional view that investigation of performance will proceed only so far as understanding of underlying competence permits. Furthermore, recent work on performance seems to give new support to this assumption. To my knowledge, the only concrete results that have been achieved and the only clear suggestions that have been put forth concerning the theory of performance . . . have come from studies of performance models that

¹Noam Chomsky, Aspects of the Theory of Syntax, p. 9.

incorporate generative grammars of specific kinds-- that is, from studies that have been based on assumptions about underlying competence.²

Later, in a section on "Linguistic Theory and Language Learning," while comparing empirical taxonomic grammars and generative theories, he writes:

. . . A general linguistic theory of the sort roughly described earlier, and elaborated in more detail in the following chapters and in other studies of transformational grammar, must therefore be regarded as a specific hypothesis, of an essentially rationalist cast, as to the nature of mental structures and processes³

In the light of such statements, it is difficult to say whether Chomsky believes in the psychological reality of the transformational process or not. The model which he presents in Aspects of the Theory of Syntax⁴ follows a schematic illustrated in Figure III of this study, with semantics entering the "flow chart" at the time of lexical insertion. But surely such a model, if taken to represent a model of performance, is absurd. At the phrase structure level of such a grammar, only reference to semantics could govern the selection of various optional elements in the P.S. derivation. Chomsky himself says in a note,

. . . I see no plausibility at all to the assumption that the speaker must uniformly select sentence type, then determine subcategories, etc., finally at the last stage, deciding what he is going to talk about; . . .⁵

²Chomsky, *ibid.*, p. 10.

³Chomsky, *ibid.*, p. 53.

⁴*Op. cit.*

⁵Chomsky, *ibid.*, p. 197.

Consequently, it is somewhat difficult to determine precisely whether or not Chomsky does credit his model with some correspondence to processes of performance or production. One gets the impression that he would support the concepts of deep structure, surface structure, and transformational processes to connect them as being elements of the production process.

Because one cannot reduce mental processes to empirical evidence, the existence or non-existence of any feature of the language production mechanism can only be approached by means of inference from a corpus of what has been said, or what subjects will say under test conditions. Such inference from product to process must, of course, remain suspect. Further, because language is both semantics (bundles of meaning) and syntax (structured strings of morphemes or phonemes), it is easy to attribute experimental effects to one which are the result of, or which are shared with, the other. Figure V illustrates one possible model of language production in very simplified form. Reference to one's own "thinking and speaking" processes supports several of the postulates of this schema: for example, it seems logical that the semantic component, under the influence of an intent to communicate, must not only begin the process of language production by supplying "semantic primitives" or "notions" of some kind, but must also monitor all steps in the entire process in order to verify that what was intended in the way of communicative production is, in fact, happening. Note also that the syntactic and transformational components are sources

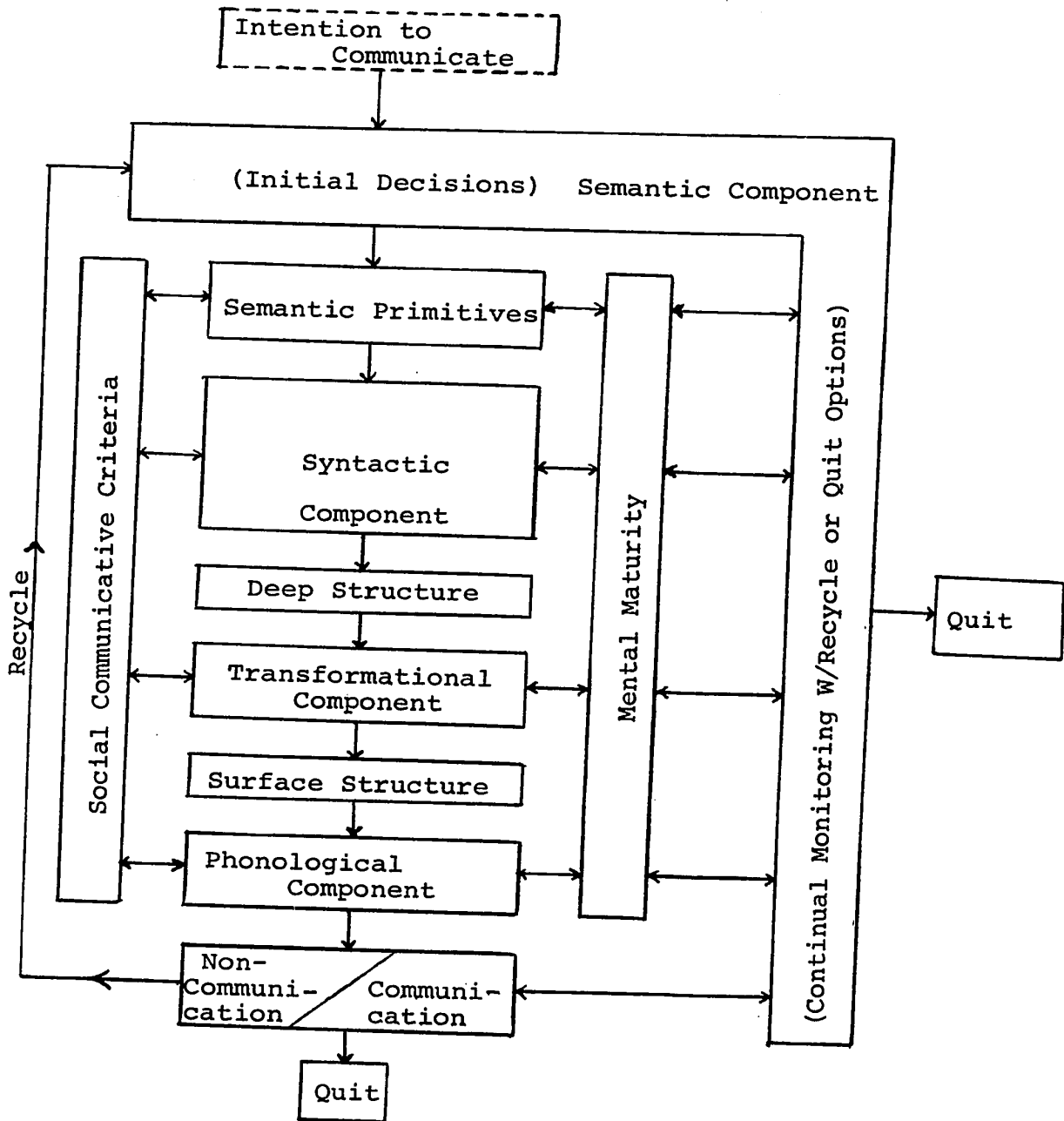


FIGURE V
SUGGESTED SPEECH PRODUCTION MODEL

of general and somewhat narrower constraints; further, the phonological component also contributes its limitations to the communicative process. The source of all of these constraints is the essentially social nature of language. To be communicable, language must be shared substantially by both the sender and receiver of a message. Consequently, both the syntactic and phonological components act as "filters" which serve to convert raw semantic primitives into units of communication. The product of the Syntactic Base/Lexicon component, the "deep structure," could be thought of as a "first approximation" of the message. The transformational component's function is to perform optional stylistic variations and to assure the grammaticality of such things as inflections. We think of involved processes such as relativization when we think of transformations within a formal written "competence" transformational grammar; in the production model being presented, such major structures would be created within the major syntactic mapping process. True, this visualization of the "phrase structure" portion of the production process makes it highly involved and complicated. However, there is no necessity in a model of mental processes to expect simplicity and generality.

To illustrate somewhat simply the reality of the reduced function of transformations envisaged in the above model, consider for a moment the difference between the often rapid outpourings of oral language as opposed to written language. The experience of the investigator is that deletions, word additions, word movements, and so forth are

common to the writer, but almost impossible for the speaker. Usually, such changes are socially motivated for stylistic or semantic purposes. Note, however, that the following strings or groups are essentially the same in semantic content:

1. The man ran. He was tired.
2. The tired man ran.
3. The running man was tired.
4. Tired, the man ran.

Sentences 2 to 4 capture the meaning of sentence 1 and while not absolute paraphrases of one another, because of slight changes in focus, they are essentially similar in content. It is reasonable that the two sentences on line 1 are the approximate form of "linguistically primitive notions" that become one of 2-4. The question is, does this combining occur at the moment when we begin to process these ideas syntactically, or might the short sentences of line 1 undergo transformation later?

To summarize, the investigator accepts the possibility of some level of production which approximates the "deep structure" level, although the extent of the transformational processes separating such deep structures is somewhat unclear. It might even be that the majority of transformational processes postulated by writers of formal grammars are performed within more complicated levels of initial synthesis, leaving very little or nothing for a later separate transformational level of process to perform.

The following studies are directed at answering questions concerning the manner in which people analyze and use language. The method and design of this study would gain strength if the following could be put on some empirical base:

1. the reality of language users' breaking sentences into syntactic units or structures,
2. evidence that people use or process these structures with varying degrees of ease,
3. indications that syntax influences semantics in some direct fashion and assists comprehension,
4. the existence of separate deep and surface structures of language, and
5. the existence of transformational processes linking these two levels.

Literature dealing with the first of these problems will be examined under the heading of "segmentation," 2 and 3 will be dealt with under "syntax, semantics and comprehension," and the last two reviews will be headed "deep and surface structure" and "transformations."

A. SEGMENTATION

The question of whether or not people break up language into units which coincide with syntactic structures was examined for the first time in any systematic way by Fodor and Bever,⁶ who had subjects listen to recorded sentences.

⁶J. A. Fodor and T. G. Bever, "The Psychological Reality of Linguistic Segments," Journal of Verbal Learning and Verbal Behavior, IV, (1965), pp. 414-420.

Superimposed on the sentences were clicks and background noise. When recalling what they had heard, subjects tended to keep syntactic units intact, and "remember" the clicks as occurring between units of structure. A later experiment by Garrett, Bever, and Fodor using a similar technique had comparable results.⁷

Graf and Torrey⁸ broke written material into segments and presented it to subjects using a Craig Reader, a device which shows only one line at a time. Presentation time was somewhat too fast for full comprehension. The story in which breaks were made corresponding to syntactic units was remembered much better than passages broken randomly.

A somewhat different approach used by Suci, Ammon, and Gamlin⁹ was termed the "probe" method. Subjects heard a sentence followed by a single word from that sentence (the "probe"). They were asked to recall and repeat the word

⁷M. Garrett, T. A. Bever and J. A. Fodor, "The Active Use of Grammar in Speech Perception," Perception and Psycholinguistics, I (1966), pp. 30-32.

⁸Richard Graf and Jane W. Torrey, "Perception of Phrase Structure in Written Language," Proceedings of the Seventy-Fourth Annual Convention of the American Psychological Association. Washington: American Psychological Association Inc., 1966, pp. 83-84.

⁹George J. Suci, Paul Ammon, and Peter Gamlin, "The Validity of the Probe-Latency Technique for Assessing Structure in Language," Language and Speed, X (1967), pp. 69-80.

following the "probe." Latencies of correct responses were greatest across major syntactic divisions, supporting the theory that people code language into major syntactic units.

The most recent study, by Bever, Lackner and Kirk,¹⁰ returned to the device of subjects subjectively locating clicks within heard language structures. A more sophisticated analysis of language structures was used in the analysis, however. Results with undergraduate subjects indicated that within-clause phrase structure boundaries do not significantly affect the segmentation of spoken sentences. Further, divisions between underlying structures of sentences determine segmentation even in the absence of corresponding explicit clause divisions in surface phrase structure. They claim their results support a model of speech processing in which the listener actively segments and organizes spoken sequences into potential syntactic structures.

The ideas concerning the manner in which language users "break up" or "segment" language are reminiscent of Miller's "chunking hypothesis,"¹¹ which states that people have an upper limit of seven (plus or minus two) on the number of "bits of information" they can deal with at any one time, and that the way we expand our capacity is by

¹⁰T. G. Bever, J. R. Lackner, and R. Kirk, "The Underlying Structure of Sentences are the Primary Units of Immediate Speech Processing," Perception and Psychophysics, 1969 (4), pp. 225-234.

¹¹G. A. Miller, "The Magic Number of Seven, Plus or Minus Two," Psychological Review, LXIII (1956), pp. 81-97.

"chunking"; that is, increasing the size of our bits. In language, this would mean dealing with larger syntactic units instead of individual words. Children would expand the size of their "chunks," reaching a particular level or plateau of "chunk size" which suited them.

B. SYNTAX, SEMANTICS, AND COMPREHENSION

One quite early study of the effect of syntactic ordering on comprehension was conducted by Miller, Heise and Lichten.¹² Their method, which became quite common in studies on comprehension, was to ask subjects to repeat something they had heard. Results showed that words are perceived more accurately if heard in sentence context than if heard merely as a string of words. The investigators believed that levels of expectation and anticipation and other context clues help the listener; but they left open the question of whether or not they were observing nothing more than an elaborate conditioning effect.

A fairly complete theory of the manner in which syntactic complexity causes semantic and comprehension problems was suggested by Yngve,¹³ using an "immediate

¹²George A. Miller, George Heise, and William Lichten, "The Intelligibility of Speech as a Function of the Context of the Test Materials," Journal of Experimental Psychology, XLI (1951), pp. 329-335.

¹³Y. Yngve, "A Model and a Hypothesis for Language Structure," Proceedings of the American Philosophical Society, 140 (1960), pp. 444-466.

Miller,¹⁴ objecting to the equation of meaning with conditioning, felt that something else, "something that linguists usually call 'grammar',"¹⁵ was involved. He felt that there were relationships within sentences, and that some words within sentences were more closely linked than others. He asks, "Is there any solid empirical evidence for the psychological reality of syntactic categories?"¹⁶ He cites two studies: in the first, Ervin¹⁷ found that the responses of adult subjects to a word association test tended to be the same syntactic categories as the stimulus words. The second study, by Glanzer,¹⁸ involved word association tasks and paired associates. Glanzer found that adult subjects found it easier to learn associations between nonsense words and form (or content) words, such as nouns, verbs, adjectives,

¹⁴George A. Miller, "Some Psychological Studies of Grammar," American Psychologist, XVII (1962), pp. 748-762

¹⁵Miller, *ibid.*, p. 748.

¹⁶Miller, *ibid.*, p. 750.

¹⁷S. M. Ervin, "Changes With Age in Verbal Determinants of Word Association," American Journal of Psychology, LXXIV (1961), pp. 361-372.

¹⁸M. Glanzer, "Grammatical Category: A Rote Learning and Word Association Analysis," Journal of Verbal Learning and Verbal Behavior, 1 (1962), pp. 31-41.

etc. than between nonsense words and function words (prepositions, conjunctions, etc.). For example, "jig-food" would be more easily learned than "of-tah." However, when function words were placed in positions where one might expect them--linking, joining, and so forth,--that is, in positions approximating their function in English, the function words were more easily learned. For example, "kex and wog" would be learned more easily than "yig-food-seb."

Epstein^{19, 20, 21} conducted a series of studies investigating whether it was possible to separate syntactic and semantic elements using nonsense words. His findings supported Miller's "chunking hypothesis."²² In a later experiment, Epstein investigated the differences between structured and unstructured materials, his criterion being temporal order. Subjects were required to learn both structured and unstructured material both forward and backwards. The structured material

¹⁹William Epstein, "The Influence of Syntactical Structure on Learning," American Journal of Psychology, LXXIV (1961), p.81.

²⁰William Epstein, "A Further Study of the Influence of Syntactical Structure on Learning," American Journal of Psychology, LXXV (1962), pp. 121-126.

²¹William Epstein, "Temporal Schemata in Syntactically Structured Material," Journal of General Psychology, LCVIII (1963), pp. 157-164.

²²Miller, op. cit.

(that is, structured into syntactic patterns of English) were learned much more quickly forwards; the unstructured materials were learned somewhat more quickly backwards.

An original approach was used by Miller and Isard,²³ whose subjects heard some 150 passages on a tape with background noise "masking" what they heard. Fifty of the passages were five-word grammatical sentences, fifty were anomalous, and fifty were five-word ungrammatical sentences. The subjects, undergraduates, were asked to repeat aloud each passage after hearing it. Scoring was on the basis of principal words remembered and number of complete sentences. Results showed scores of 88.6% for grammatical sentences, 79.3% for anomalous sentences, and 56.1% for ungrammatical sentences, the differences all being significant. They concluded that syntactic structures do take a part in subjects' ability to hear and repeat sentences correctly. However, they also felt that semantics must play a considerable part, as indicated by the results using anomalous sentences.

The method of Marks and Miller²⁴ involved the presentation of lists of five-word groups or strings to subjects for memorization. The first group consisted of five normal sentences, the word order being adjective--plural noun--verb--adjective--plural noun. Group two was prepared taking the

²³George A. Miller and Stephen Isard, "Some Perceptual Consequences of Linguistic Rules," Journal of Verbal Learning and Verbal Behavior, II (1963), pp. 217-228.

²⁴Lawrence E. Marks and George A. Miller, "The Role of Semantic and Syntactic Constraints in the Memorization of English Sentences," Journal of Verbal Learning and Verbal Behavior, III (1964), pp. 1-5.

first word from the first sentence of group one, the second word from sentence two, and so on. Group three had five more strings rearranged anomalously from group one. Finally, group four consisted of five-word scrambled lists. Learning proved easiest for grammatical sentences, hardest for word lists. Errors were divided into those judged semantic (moving words from one string to another) and syntactic (bound-morpheme errors and inversions). Semantic errors were more common in anomalous strings and word lists; syntactic errors occurred most frequently in anagram strings and word lists where syntactic rules were most violated. They concluded, ". . . these errors give support to the contention that syntactic and semantic rules have psychological as well as linguistic reality."²⁵

A different approach was used by Schlislinger,²⁶ who asked his adult subjects to read materials from marked passages. Word length was carefully controlled. At intervals, the light was switched off, and subjects were asked to repeat as much as they could remember of what they had read. Results showed that the last word seen and recalled was usually the end of a syntactic element. A second experiment which carefully controlled for sentence length had very similar results. The same results were obtained from both slow and fast readers.

²⁵Marks and Miller, *ibid.*, p. 4.

²⁶I. M. Schlislinger, "The Influence of Sentence Structure on the Reading Process," Technical Report, Number 24, United States Office of Naval Research, Informations Systems Branch (Jerusalem, Israel) October, 1966.

Using similar presentation methods, Forster²⁷ investigated whether the effect of syntactic structures aiding learning would disappear if subjects were not required to learn items or facts in any particular order. Results bore out the thesis that syntactic structuring of material facilitates recall.

Epstein²⁸ used an approach similar to Marks and Miller²⁹ in that he prepared structured lists of materials. He wished to investigate the effect of syntactic structures other than active verb forms. Anomalous strings were constructed, as well as strings containing active verbs, passive verbs, and nominalization forms. Not only did he find that all of the structured forms facilitated learning, but he also found a clear variation in the difficulty of the selected structures, the number of words learned being greatest for active verb forms, intermediate for passives, and least for nominalizations. Such a finding is of considerable interest to the present study, and further reference to this study will be made, both in the section dealing with transformations, and in the summary.

²⁷Kenneth I. Forster, "The Effect of Syntactic Structures on Nonordered Recall," Journal of Verbal Learning and Verbal Behavior, V (1966), pp. 292-297.

²⁸William Epstein, "Some Conditions of the Influence of Syntactical Structure on Learning: Grammatical Transformation, Learning Instructions, and 'Chunking'," Journal of Verbal Learning and Verbal Behavior, VI (1967), pp. 415-419.

²⁹Marks and Miller, op. cit.

In a large study involving 1,016 children of ages from 5 to 9, Herriot³⁰ asked his subjects to match pictures to active and passive sentences, some of which had nonsense words substituted for some, all, or none of their content words. Significant differences in levels of comprehension were found when nonsense words were inserted, although older children succeeded in correctly matching picture and passage better than younger children. Wrote Herriot, "It may be concluded that the syntax of the sentence has semantic reference."³¹

A recent study by Perfetti and Goodman³² returned to Yngve's model of "depth"³³ in order to explore its viability in some depth. In their first experiment, forty subjects were presented, first with 20 sentences of "depth 7," a complicated structure according to Yngve's model and then were presented with a sentence of the same semantic content (meaning) constructed to be of "shallow" depth. Findings indicated that the location of embeddings was more important to subjects' comprehension than depth. In a second experiment, the difficulty of adverbial noun phrases and adjectival noun phrases was compared.

³⁰Peter Herriot, "The Comprehension of Syntax," Child Development, 39 (March, 1968), pp. 273-282.

³¹Herriot, *ibid.*, p. 281.

³²Charles A. Perfetti and Doba Goodman, "Memory for Sentences and Noun Phrases of Extreme Depth," Quarterly Journal of Experimental Psychology, XXIII (1971), pp. 22-33.

³³Yngve, *op. cit.*

Twenty-one sentences containing three types of structures were presented aurally to subjects, and after a ten-second delay, they were asked to reproduce what they had heard. Little difference resulted, with the position of keywords seemingly more important than total structure. Their conclusion was that Yngve's model was not useful and that the type of "depth" he used as a predictor of difficulty was not valid from their findings.

C. DEEP AND SURFACE STRUCTURE

Using Chomsky's model of grammar, Mehler³⁴ constructed groups of eight sentences incorporating the following types: kernels, negatives, passives, questions, negative-questions, negative-passives, question-passives, and negative-question-passives. Eighty undergraduate subjects were exposed to groups of sentences, and then asked to recall as many as possible. Single prompt words were given. Mehler discovered that in many cases, the "gist" of the sentence might be there while the exact transformed structure would be gone. He concluded that his subjects were storing "kernels" and transformations (or what he termed "syntactic corrections") separately. He did not speculate as to how these elements were encoded or learned. In a later experiment with Carey,³⁵

³⁴Jacques Mehler, "Some Effects of Grammatical Transformations on the Recall of English Sentences," Journal of Verbal Learning and Verbal Behavior, II (1963) pp. 346-351.

³⁵Jacques Mehler and Peter Carey, "The Role of Surface Structure and Base Structure in the Perception of Sentences," Journal of Verbal Learning and Verbal Behavior, VI (1967), pp. 335-338.

Mehler continued to examine the reality of deep and surface structures. The investigators produced a series of sentences, all beginning with "They are . . . ," but incorporating four syntactic structures. For example, sentences such as "They are conflicting desires," and "They are forecasting cyclones," were presented to subjects along with background noise, and the relative accuracy of perception or recall was recorded. When a sentence differing in surface structure was presented following ten sentences of uniform structure, errors in perception increased significantly. Similar results were obtained when sentences of similar surface structure but different deep structure were presented. Guessing on the part of the subjects was encouraged, and it was noted that they often followed syntactic, rather than phonological clues in substituting words. They concluded that their results clearly indicated separate levels of language processing.

A study of very similar nature to the above was conducted by Rohrman³⁶ using 352 undergraduates in five separate experiments. The purpose of the study was to determine whether surface or deep structure more adequately characterizes the memory representative of a sentence. Experiments 1-4 showed that with materials of identical

³⁶Nicholas L. Rohrman, "The Role of Syntactic Structure in the Recall of English Nominalizations," Journal of Verbal Learning and Verbal Behavior, VII (1968), pp. 904-912.

surface structure, but differing in deep structure complexity, differences in recall latencies are predictable from the latter. Rohrman concluded that underlying structure is what is stored in the memory. Experiment 5 determined whether deep structure complexity (as measured by counting nodes) or transformational history was more important in determining recall. Results favored deep structure complexity over transformational history.

Another experiment along similar lines was conducted by Davidson³⁷ using two groups of undergraduates who learned two lists of sentences which differed in deep structures, but which were identical in surface structures. Word-to-word transition errors were used as a scoring device, and frequency of errors showed quite different recall patterns, corresponding to different deep structure patterns.

D. TRANSFORMATIONS

Probably the earliest research into the reality of the transformational process was conducted by Miller, McKean and Slobin.³⁸ The experiment was designed to demonstrate that the relative difficulty of making systematic changes in

³⁷Robert E. Davidson, "Transitional Errors and Deep Structure Differences," Psychonomic Science, XIV (1969), pp. 293-294.

³⁸G. A. Miller, K. McKean and D. Slobin, "The Exploration of Transformations by Sentence Matching," in G. A. Miller, "Some Psychological Studies of Grammar," American Psychologist, XVII (1962), pp. 748-762.

sentences could be predicted as a function of transformational complexity of the sentences involved. Paired sentences differing by only one or two transformations were randomly distributed within their respective groups, and subjects were required to match pairs. Subjects were instructed as to the nature of the transformations involved. A base search time was determined by having subjects match two simple active affirmative declarative sentences in random groups, and, for example, two passives. Results from the experimental tasks gave the following processing times: (Note--SAAD means "simple, active, affirmative, declarative, or "kernel.")

| | | | |
|------|---|----------|-------------------------------|
| SAAD | - | N | 1.1 Secs. beyond search time, |
| SAAD | - | PASS | 1.4 |
| PASS | - | PASS NEG | 1.7 |
| NEG | - | PASS NEG | 1.9 |
| SAAD | - | PASS NEG | 2.7 |
| NEG | - | PASS | 3.5 |

These figures support the hypothesis that processing time is a function of transformational history, thus providing some evidence to support the Derivational Theory of Complexity (D.T.C.) as well. Mehler's 1963 study,³⁹ besides supporting separate levels of processing, gave further evidence to support D.T.C.

McMahon's⁴⁰ approach required subjects to signal whether a statement they were looking at (for example,

³⁹Mehler, op. cit.

⁴⁰E. McMahon, "Grammatical Analysis as Part of Understanding," (Unpublished Doctoral dissertation, Harvard University) 1963.

"15 precedes 13," or "3 is preceded by 7") was true or false by depressing appropriate keys attached to timers. He found negative sentences to take longer to process than actives or passives. Summarized, his results were as follows:

| | | |
|---------------------|---|------------------|
| Active Affirmative | - | small difference |
| Passive Affirmative | - | large difference |
| Active Negative | - | small difference |
| Passive Negative | | |

One might interpret such results on a semantic criterion: the transformation whose effect upon truth value is most difficult to determine was easiest, or, put another way, a transformation requiring something beyond acceptance or affirmation was more difficult. These results compare interestingly with those of Miller and McKean,⁴¹ who wished to improve upon the techniques of the earlier study of Miller, McKean and Slobin.⁴² Presentation of sentences was singly by means of a tachiscope. When the subject had performed the required transformation, he pressed a button which stopped a timer and presented a search list. Results were more uniform than in the earlier study. They found that in all cases, single transformation (for example, SAAD→PASS.) was easier than double transformation (SAAD→PASS./NEG).

⁴¹G. A. Miller and K. McKean, "A Chronometric Study of Some Relations Between Sentences," Quarterly Journal of Experimental Psychology, XVI (1954), pp. 297-308.

⁴²Miller, McKean and Slobin, op. cit.

SAAD transformations were easiest, followed by negatives, and passives were more difficult. Double transformations of any SAAD form were easier than ones involving transforms of either NEG. or PASS. sentences. The difference in order of these findings and those of McMahon⁴³ may be the result of the truth-value judgments required in the earlier study.

Two studies by Coleman⁴⁴ and Coleman and Blumenfeld⁴⁵ in this area are of special interest to the present study, as they used the "Cloze" procedure to determine relative difficulties of structures incorporating active verbs and their corresponding nominalizations. Results significantly favored the active verb forms. Coleman also hypothesized that subjects would "de-transform" passives into kernels, a prediction similar to that of Mehler⁴⁶ and Miller.⁴⁷ Coleman had his subjects match active and passive forms. His findings showed 56 passives retained as actives, and only 33 actives retained as passives, a result which he believed supported the concept of re-coding of sentences as kernels.

⁴³McMahon, op. cit.

⁴⁴E. B. Coleman, "Learning of Prose Written in Four Grammatical Transformations," Journal of American Psychology, XLIX (1965), pp. 332-341.

⁴⁵E. B. Coleman and J. P. Blumenfeld, "Cloze Scores of Nominalizations and their Grammatical Transformations," Psychological Reports, XIII (1963), pp. 651-654.

⁴⁶Mehler, op. cit.

⁴⁷Miller, op. cit.

Another study by Savin and Perchonock⁴⁸ explored short-span memory capacity, a factor vital to studies such as that of Coleman and Blumenfeld.⁴⁹ Savin and Perchonock pictured short-span memory having a fixed capacity. They presented sentences involving passives, negatives, passive-negatives, and so forth. Each sentence was followed immediately by a string of eight random words. Subjects were asked to recall immediately the entire sentence, and then as many of the "extra" words as they could. Results in terms of the number of words remembered supported the thesis that sentences involving more than one transformation used up more "memory capacity." A second part of this study investigated the number of random words recalled from within a "frame" of sentences of different structures. Surrounding "frames" were made up of: kernel to passive; emphatic to emphatic-passive; negative to negative-passive; question to question-passive; and question-negative to question-negative-passive. The number of words remembered in each interval proved approximately equal, supporting the hypothesis that transformations of each of the surrounding sentences are remembered independently and without regard to the rest of the sentence. Such findings must bring doubts to bear upon

⁴⁸Harris B. Savin and Ellen Perchonock, "Grammatical Structure and the Immediate Recall of English Sentences," Journal of Verbal Learning and Verbal Behavior, IV (1965), pp. 348-353.

⁴⁹Coleman and Blumenfeld, op. cit.

the concept of "re-coding," but seem to agree with the "total number of transformations" ideas of the D.T.C.

Gough⁵⁰ and Slobin⁵¹ both used pictures which subjects of various ages were asked to differentiate in order to verify a grammatical structure they had just heard. The pictures were presented immediately after the sentence was heard. Slobin's subjects--from kindergarten, grades 1, 4, 6, and adults--found kernels, negatives, passives, and negative-passives to be increasingly difficult in that order at all age levels.

Gough reasoned that if a delay was provided after presentation of grammatical structures, and verification then sought, the verification times alone would not differ for different structures. That is, he reasoned that subjects would do their "grammatical analysis" during the delay and verification should consist of simple yes/no decisions. He found that subjects still took longer to verify passives and negatives than actives. His conclusion was that the time differences displayed in his study (and, by inference, in Slobin's) were functions of both the process of verification

⁵⁰ Phillip B. Gough, "The Verification of Sentences: The Effect of Delay of Evidence and Sentence Length," Journal of Verbal Learning and Verbal Behavior, V (1966), pp. 492-496.

⁵¹ Dan I. Slobin, "Grammatical Transformations and Sentence Completion in Childhood and Adulthood," Journal of Verbal Learning and Verbal Behavior, V (1966), pp. 219-227.

(a process which McMahon⁵² felt involved semantic judgments) and the processes of breaking down structures. Consequently, while Gough himself felt his study did not confirm the D.T.C., his results do not prove conclusively that the total time differences are attributable to verification, rather than comprehension.

In 1967, Fodor and Garrett,⁵³ reflecting upon a number of studies conducted up to that time, and especially those of Miller, formalized the Derivational Theory of Complexity as follows:

D.T.C. can be made explicit in the following way. Consider a generative grammar G of the language L and a sentence S in the range of G . It is possible in principle to define a metric which, for every pair (G, S) specifies the number N_1 of rules (or elementary operations or whatever) G_1 requires to generate S_1 . D.T.C. in its strongest form is the claim that the size of N_1 is an index of the complexity of S_1 . In particular, two sentences assigned the same number are equally complex, and, of two sentences assigned different numbers, the larger number is assigned to the more complex sentence.⁵⁴

Note that in its stated form, the D.T.C. makes no specific reference to language production; it is a very simple metric defined in terms which are not precise, but which are apparently open to interpretation. Restated, it says

⁵²McMahon, op. cit.

⁵³J. A. Fodor and M. Garrett, "Some Syntactic Determinants of Sentential Complexity," Perception and Psychophysics, II (July, 1967), pp. 289-296.

⁵⁴Fodor and Garrett, *ibid.*, p. 289.

something like, "Equals are equal; and given two different sized things, the larger is the bigger." Applied to sentences, the D.T.C. would enjoin an investigator to choose a metric (transformational operations, elementary operations, number of clauses, ratio of subordinate clauses to principal clauses, number of hard words, total number of words, and so on) and begin counting in order to compare the difficulty of any two sentences. That the D.T.C. has come to be interpreted as referring to transformational rules is simply a result of the context within which this theory was stated. Upon examination, however, there are two interpretations which are commonly placed upon the D.T.C. in investigations relating it to actual language performance. One interpretation says we must base our metric of "sentence difficulty" upon a simple counting of transformational steps (that is, a passivized sentence has been derived by means of one operation, a negative-passive by means of two). A second interpretation would require a determination of relative difficulty of specific transformations by means of counting elementary operations or steps within each transformation, and then a counting of appropriately weighted values within sentences. Thus, present discussion of the viability of Fodor and Garrett's metric concerns itself with whether the kind of transformation is as important as the total number of transformations. Intuition (at best, a poor research instrument) would hold that kind of structural

change performed is certainly a factor. For example, it seems reasonable that a simple deletion would be less difficult than a transposition of elements.

Investigations of this theory have had somewhat mixed results. Salinger and Eckerman,⁵⁵ returning to methods used earlier by Mehler,⁵⁶ had subjects listen to strings, and then write down as much as they could remember. They controlled for a number of factors ignored in earlier studies, such as sentence length, order of words, and number and method of presentation. Salinger and Eckerman, referring to earlier studies, chose simple-active-affirmative and passive-negative-questions as their experimental structures in order to increase the possibility of getting clear differences in response. Precisely how they managed to make simple sentences as long as passive-negative-questions without introducing further factors was not made clear. Findings indicated a difference on first presentation, but little difference on second or subsequent presentations, a possible result of learning effects. They state that their findings were inconclusive as to the relative difficulty of the selected structures. However, their results seem to support that a greater number of transformations increases

⁵⁵Kurt Salinger and Carol Eckerman, "Grammar and the Recall of Chains of Verbal Responses," Journal of Verbal Learning and Verbal Behavior, VI (1967), pp. 232-239.

⁵⁶Mehler, op. cit.

"difficulty." Essentially identical results were obtained by Compton.⁵⁷ A similar study by Epstein,⁵⁸ already reviewed in another connection, found the following order of difficulty of structures, ranging from easiest to hardest: active verb forms, passive verb forms, and nominalizations.

An experiment claimed to support the "transformational decoding hypothesis" was conducted by Morris, Rakine and Reber.⁵⁹ While their design was very much like those of Gough⁶⁰ and Slobin,⁶¹ their response device which used a system of keys for response was somewhat more sophisticated. They measured comprehension of sentences incorporating active, passive, negative, and negative-passive forms by subjects who reconstructed sentences by manipulating keys. Latency and error analysis measures were applied for both immediate responses and responses delayed 3 seconds. Results duplicated those of the two studies cited earlier in almost all respects.

⁵⁷ Arthur J. Compton, "Aural Presentation of Different Syntactic Structures and Lengths," Language and Speech, X (April, 1967), pp. 81-87.

⁵⁸ Epstein, 1967, op. cit.

⁵⁹ Val A. Morris, Fred C. Rakine and Arthur S. Reber, "Sentence Comprehension, Grammatical Transformations, and Response Availability," Journal of Verbal Learning and Verbal Behavior, VI (1968), 1113-1115.

⁶⁰ Gough, op. cit.

⁶¹ Slobin, op. cit.

A study by Ammon⁶² using "probe" technique identified four factors which he felt influenced ability to deal with sentences at the level of surface structure. These features were: separation of related elements, presence of relative as opposed to personal pronoun forms, confusability of elements, and congruence of elements. Ammon also used a series of questions regarding relationships of ideas within sentences. His conclusion was that a number of elements can contribute to sentence complexity, which in turn increases latency of subjects' responses.

One experiment which, the investigator concluded, tended to show that subjects do not transform sentences automatically after hearing them was conducted by Wright.⁶³ Working with only active and passive forms, she asked 80 subjects either an active or a passive question five seconds after they heard a simple active statement. The fewer errors when the question and statement were in the same voice indicated to Wright that her subjects did not "de-transform" passives into actives as a matter of course. The present investigator, however, would suggest that the passive form, with its transposition of N.P. elements, introduces elements of focus and emphasis which, in many ways negate comparisons

⁶²Paul R. Ammon, "Some Research on the Comprehension of Sentences," Child Study, XXX (Winter, 1968-69), pp. 14-18.

⁶³Patricia Wright, "Transformations and the Understanding of Sentences," Language and Speech, III (1969), pp. 156-166.

of active and passive forms.

An interesting experiment by Levelt and Ouwenal⁶⁴ used 48 French sentences (with ambiguous surface structures) both in isolation and in context. Subjects were confused as to meaning far more often when sentences were in isolation, indicating that transformational processes seem to operate automatically unless there is context which semantically blocks one or another possible reading. Such a finding also seems to support the skeleton production model put forward in the introduction to this section wherein there is constant feedback between all elements of speech production and a semantic component.

A study which examined the relative difficulty of active and passive forms, as well as reading difficulty analysis, was conducted by DeVito.⁶⁵ Twelve sentences, half active and half passive, were spoken to 25 undergraduate subjects, who were given a prompt before trying to repeat what they heard. Significant differences were found in comparing frequency of accurate response. In a second part of the study, parallel active and passive passages were constructed and analyzed according to three widely used

⁶⁴W. J. Levelt and G. R. Ouwenal, "The Perception of French Sentences With a Surface Structure of Ambiguity," Nederlands Tijdschrift voor de Psychologie en haar Grenogebieden, XXIV (1969), pp. 245-248.

⁶⁵Joseph A. DeVito, "Some Psycholinguistic Aspects of Active and Passive Sentences," Quarterly Journal of Speech, LV (December, 1969), pp. 401-406.

"readability" formulas; the Flesch,⁶⁶ the Farr-Jenkins-Paterson,⁶⁷ and the Gunning "Fog Index."⁶⁸ The Flesch and Farr-Jenkins-Paterson formulas showed the passive to be slightly more readable, while the Gunning Fog showed the active to be somewhat easier to read. In part 3, DeVito constructed "Cloze" tests on both passages from part 2, and scored them using both verbatim and form-class methods. The verbatim method indicated the passive form to be easier, possibly because of the high degree of redundancy involved. Form class scoring favored the active form. These results indicate some of the difficulties which the reading formula approach to the complexity of sentences cannot resolve. As already pointed out, the enormous redundancy of any passage limited to passive verbs only would destroy the effectiveness of any "Cloze" test.

Of great interest to the present study is that of Fagan,⁶⁹ who was working within the framework of reading difficulty and sentence complexity. He was concerned with the difficulty of the structures of language appearing in basal readers at the grades 3, 4 and 5 levels. First, 21

⁶⁶Flesch, op. cit.

⁶⁷James N. Farr, James J. Jenkins, and Donald G. Paterson, "Simplification of the Flesch Reading Ease Formula," Journal of Applied Psychology, XXXV (October, 1951), pp. 333-337.

⁶⁸Gunning, op. cit.

⁶⁹Fagan, op. cit.

passages were chosen from readers, and a transformational grammar capable of generating the structures present was prepared. These passages were tested for readability using the "Cloze" procedure. The second part of the study involved the re-writing of three randomly chosen stories in each. Among the types of transformations used were: Embeddings, Conjoinings, Deletions, so-called "Simple" transformations, and others. "Cloze" procedure was applied. Part three consisted of adding sentences from one of the classes above to passages drawn from the same books as were used above, and applying "Cloze" again. Three methods were used in scoring "Cloze" tests: a basic method requiring verbatim responses, a method which determined difficulty indices for passages, sentences, and transformational units, and a third method investigated the grammaticality of the form which students inserted. Results indicated that the presence of Embedding and Deletion transformations contributed to passage and sentence difficulty. The relative difficulty of Deletion transformations also had an effect upon difficulty. The type of Simple and Conjoining transformations was more important than mere presence in determining difficulty. On the whole, Conjoining transformations were easiest. It appeared that sentence difficulty was more dependent upon the presence and difficulty of transformations than was passage difficulty, a result which seems

to bear out the findings of Levelt and Ouwenal.⁷⁰ It also appeared that the presence of structures commonly found in oral language--vocatives, expletives, and so on--made passages easier. The number of transforms per sentence did not seem to contribute to sentence difficulty, nor did the total number of words per sentence. "Cloze" scores correlated significantly with age, sex, grade, reading achievement, mental ability, and socio-economic status. Embedding transformations were found to be most common in occurrence within the basal reader materials used, followed by Simple, Conjoining, Position, Shift, and Deletion. Results across grades were consistent with respect to the difficulty of structures. Fagan felt that the D.T.C. was not supported on the basis of a simple count of transformations; however, a count of elementary operations resulted in findings termed "controversial."⁷¹ Several things set this study apart. The use of "Cloze" procedures adapted to determine the difficulty of certain specific structures is noteworthy, and does away with problems which plague other scoring methods (latency periods cannot be divided into "solving" and "verification" factors; short-term memory is, at best, a highly complex and variable thing.) The use of a wider variety of structure types along lines suggested by

⁷⁰Op. cit.

⁷¹Fagan, *ibid.*, p. vi.

Bateman and Zidonis⁷² is an important expansion of theory. Note that the structures used in most of the earlier studies (passive, negative, question) are all grouped into "Simple" transforms. Also, Fagan went beyond sentences in isolation (the usual method) to determine the difficulty of entire passages containing certain structures. His findings, that type of transformation has more to do with difficulty than sheer numbers of such structures, is a departure, and seems to indicate, at least to some extent, the fallacies of the "count the elements" approach, no matter what those elements might be. Several of Fagan's findings will be compared with the results of the present investigation in the final chapter of this study.

Summary

a. Segmentation. No one would doubt that humans break up the stream of their language outpourings into units; the question is, what is the nature of those units? It would seem that we segment our speech into units which have been termed syntactic, but which in fact would correspond to semantic units as well. Further, it would seem as though our capacity to process these units has an upper bound. The suggested manner in which we can extend our ability to deal with ideas and information at any one

⁷² Donald Bateman and Frank Zidonis, The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Grades, Champaign, Illinois: National Council of Teachers of English, 1966.

time is by enlarging the size or scope of our "chunks" of information. Within language processes, this would consist of enlarging or expanding by linguistic (or grammatical) means whatever is encompassed within syntactic-semantic units.

b. Syntax, semantics, and comprehension. What we hear or get out of language is a function of many things, including our expectations or anticipations, a result of the considerable redundancy of language. But the structures of language also have their effect upon comprehension. Long involved modifiers may contribute to language difficulty; simple active verb forms make sentences easy to comprehend, while passive forms of the same sentences are somewhat more difficult. The evidence seems clear, however, that grammatical ordering of language facilitates its learning and/or recall, and that the ease with which certain structures of language are processed varies in a consistent pattern. That is, certain structures seem to be more difficult to process than others. The cause of these variations could be the manner in which we process language structures: some evidence indicates we may store sentences as "kernels" with appropriate "transformational tags," although such evidence is not conclusive.

c. Deep and surface structure. Before any reality of transformational processes taking place can be established,

there must be evidence to establish the existence of "deep" and "surface" structure levels within language; otherwise, the scope of transformations would be greatly limited. On the least specific level, the fact that people tend to be able to recall the "gist" of a notion, even when they cannot recall such facts as whether the speaker's attitude was negative or positive toward that notion, gives some viability to different levels. More conclusive evidence has been produced using 1. ambiguous structures; that is, sentences which can have several meanings, and 2. groups of sentences of identical surface structure, but differing deep structure. Sensitivity to changes in deep structure has been clearly indicated by increases in "processing time" for sentences of differing deep structure.

d. Transformations. The biggest problems in coming to grips with the existence of transformational processes and a scale of relative difficulty of such processes in human language production have been in establishing consistent terms of reference and reliable metrics. Most studies consider a transformation as a single, unitary process, and do not measure the singular operations which are included within transformations. Operational problems have been encountered in using latencies as a metric, and studies involving short-term memory are dealing with a capricious capacity. There seems to be a good deal of

support, however, for the theory that some language structures are more difficult to process than others, although the variety of structures used in most studies has been limited. The theory that a structure of language is difficult in proportion to the number of "processes" involved in its derivation within the framework of formal transformational grammar theory seems not only vacuous, but has received little empirical support. A recent approach which seems promising with respect to establishing the relative difficulty of language structure is the adaptation of "Cloze" procedure from reading theory.

General Summary of Psycholinguistics

The lack of conclusive empirical evidence as to the existence or non-existence of any theorized process is not unusual within the behavioral sciences. With respect to how we process language, there seems to be sufficient evidence to draw the following tentative conclusions:

1. Human language users, whether producing or receiving linguistic data, break up streams of linguistic data into some form of syntactic/semantic units that go beyond single words.
2. These units, which are processed with varying ease, clearly aid comprehension by their presence.
3. We process language at more than one level, but whether these levels are the result of different semantic readings of the words involved, or actual syntactic levels is not clear.
4. The processes whereby a basic notion becomes negative, a question, a passive, and soon, seem to relate consistently to one another in difficulty regardless of the nature of the underlying notion.

In the light of these conclusions, the present study is interested in trying to establish the relative difficulty of a wider range of transformationally-defined language structures. Further, this study seeks to explore whether there is a relationship between a subject's ability with language structures and his knowledge of the structures to be investigated in the first section of the investigation.

The next portion of this review of literature examines studies aimed at identifying characteristics of language use which mark the "better," more developed or more "mature" writer.

4. SYNTACTIC AND WRITING ABILITY

Introduction

A number of investigators, chiefly educators interested in developing approaches to more effective language arts curricula, have attempted to determine either

1. what qualities of writing ability develop with age, or
2. what marks the superior or better writer?

Other investigators have attempted to use some of the "findings" in the language arts area in efforts to determine whether, in fact, it is possible to get students to use language better by means of some form of direct teaching. Thus the following survey of literature is broken into two parts, one dealing with "What makes a better writer?" and the other discussing, "Can we teach how to write better?" The first portion surveys studies aimed at determining what people do with language, and how their patterns of language growth change--supposedly for the better. The second portion examines some attempts which have been made to use some of the knowledge gained from the studies of part A, with attention being focussed mainly on attempts to put transformational theories to work.

Neither of these surveys is complete: reference is given to excellent summaries compiled in the past of studies in the area of language development. With respect to the usefulness of various teaching applications of knowledge gained from language research, a survey of

studies attempting to justify all of the various approaches taken would require a book several times the size of this entire study. It must suffice to say that no "best" way of teaching composition is yet to be found, and, if ever found, would probably differ for every individual student. Consequently, because at the present time a comprehensive survey would be totally inconclusive, part B of this study's survey examines several representative approaches, and then concentrates upon transformational applications.

A. LANGUAGE ABILITY AND DEVELOPMENT

The most notable surveys of studies in the area of language behavior began with the study of Heider and Heider,¹ who were concerned with the sentences used by deaf children. Harrell's² study, itself somewhat developmental in nature, supplemented Heider and Heider's earlier survey of behavior. Developmental studies were summarized with special emphasis upon early childhood changes by McCarthy,³ Carroll,⁴ and

¹F. K. Heider and G. M. Heider, "A Comparison of Sentence Structure of Deaf and Hearing Children," Psychological Monographs, LII (1940), pp. 42-103.

²Lester E. Harrell, Jr., "A Comparison of Oral and Written Language in School-Age Children," Monographs of the Society For Research in Child Development, XXII (1957).

³Dorothea A. McCarthy, "Language Development in Children," in Manual of Child Psychology, second ed., ed. Leonard Carmichael, New York: John Wiley and Sons, 1954, pp. 492-630.

⁴John B. Carroll, "Language Development in Children," in Encyclopedia of Educational Research, ed. Chester W. Harris, New York: Macmillan, 1960, pp. 744-752.

Ervin and Miller,⁵ while studies covering older children are surveyed in the work of Strickland,⁶ Loban,^{7,8,9,10} Menyuk,¹¹ Hocker,¹² Hunt,^{13,14,15,16} Sam and Stine,¹⁷ and O'Donnell,

⁵Susan M. Ervin and W. R. Miller, "Language Development," in Child Psychology, Sixty-second Yearbook, Part I, National Society for the Study of Education, ed. Harold W. Stevenson, et. al. Chicago: U. of Chicago Press, 1963, pp. 108-143.

⁶Ruth G. Strickland, The Language of Elementary School Children: Its Relationship to the Language of Reading Textbooks and the Quality of Reading of Selected Children, Bulletin of the School of Education, Indiana University, 38:4, Bloomington: Indiana University, 1962.

⁷Walter D. Loban, Language Ability in the Middle Grades of the Elementary School, Report to the U.S. Office of Education, Contract 7287 Berkley: 1961.

⁸_____, The Language of Elementary School Children, National Council of Teachers of English Research Report No. 1, Champaign, Ill.: N.C.T.E., 1963.

⁹_____, Language Ability: Grades Seven, Eight, and Nine. Report to the U.S. Office of Education, Cooperative Research Project No. 1131. Berkeley: 1964.

¹⁰_____, Language Ability: Grades Ten, Eleven, and Twelve, Report to the U.S. Office of Education, Cooperative Research Project No. 2387. Berkeley: 1967.

¹¹Paula Menyuk, "Syntactic Rules Used by Children from Preschool Through First Grade," Child Development, 35:2 (June, 1964), pp. 533-546.

¹²Mary Elsa Hocker, "Reading Materials for Children Based on Their Language Patterns of Syntax, Vocabulary, and Interests." (Unpublished Master's thesis, University of Arizona.)

¹³Kellogg W. Hunt, Differences in Grammatical Structures Written at Three Grade Levels, the Structures to be Analyzed by Transformational Methods, Report to the U.S. Office of Education, Cooperative Research Project No. 1989. Tallahassee, Florida: 1964.

¹⁴_____, Grammatical Structures Written at Three Grade Levels, National Council of Teachers of English Research Report No. 3. Champaign, Ill.: N.C.T.E., 1965.

Griffin, and Norris.¹⁸

A number of the earlier studies such as those of Bear,¹⁹ Frogner,²⁰ and Labrant²¹ discuss problems of student writing, such as incomplete and run-on sentences. These and other early studies (e.g. Davis,²² Heider and Heider)²³ were

¹⁵Kellogg W. Hunt, Sentence Structures Used by Superior Students in Grades Four and Twelve, and By Superior Adults, Report to the U.S. Office of Education, Cooperative Research Project No. 5-0313, 1966.

¹⁶_____, Syntactic Maturity in Schoolchildren And Adults, Monographs of the Society For Research in Child Development, XXXV No. 1, Chicago: University of Chicago Press, 1970.

¹⁷N. H. Sam and E. S. Stine, Structural Analysis of the Written Composition of Intermediate Grade Children, Report to the U.S. Office of Education Cooperative Research Project No. S-057. Bethlehem, Pennsylvania, 1965.

¹⁸Roy C. O'Donnell, William J. Griffin, and Raymond C. Norris, Syntax of Kindergarten and Elementary School Children: A Transformational Analysis, National Council of Teachers of English, Research Report No. 8. Champaign, Ill.: N.C.T.E., 1968.

¹⁹Mata V. Bear, "Children's Growth in the Use of Written Language," Elementary English Review, XVI (December, 1939), pp. 312-319.

²⁰Ellen Frogner, "Problems of Sentence Structure in Pupils' Themes," English Journal, XXII (November, 1933), pp. 742-749.

²¹Lou L. La Brant, "A Study of Certain Language Developments of Children in Grades Four to Twelve, Inclusive," Genetic Psychology Monographs, XIV (November, 1933), pp. 387-497.

²²Edith A. Davis, The Development of Linguistic Skill in Twins, Singletons with Siblings, and Only Children of Age Five to Ten Years, Institute of Child Welfare Monograph Series, No. 14. Minneapolis: University of Minneapolis Press, 1937.

²³Heider and Heider, op. cit.

concerned with what came to be known as the "standard approach." As in readability studies, researchers worked out methods using counting of elements: sentence length, clause length, the number of subordinate clauses per main clause (the "subordination ratio"), kinds of subordinate clauses, and total words falling into eight (or nine) "parts of speech" became the criteria. In 1954, McCarthy wrote:

Apparently, then, sentence length is a measure which continues to show increase up until maturity. The use of the measure has been criticized by some writers and a few substitute measures have been suggested, but none seems to have superseded the mean length of sentence for a reliable, easily determined, objective, quantitative, and easily understood measure of linguistic maturity. . . .

It is interesting to note from La Brant's study, however, that length of clause remains fairly constant in Grades 4 to 12, although the subordination index, or ratio of subordinate to coordinate clauses, shows an increase. Apparently length of clause is somewhat controlled or restricted by the structure of the language, and whatever increase in sentence length occurs at higher age levels is brought about largely through the addition of more subordinate clauses.²⁴

Many teachers, in the light of such statements, strove to teach students to use a greater percentage of complex and compound-complex sentences in search of more mature writing.

One problem for investigation was that of deciding precisely what constituted a sentence, for measures of "mean" sentence length were meaningless without some operable criterion. In her study, McCarthy²⁵ tried to clarify some

²⁴McCarthy, op. cit., pp. 522-523.

²⁵Dorothea A. McCarthy, Language Development of the Preschool Child, Institute of Child Welfare Monograph Series, No. 4. Minneapolis: University of Minneapolis Press, 1930.

of the difficulties in classification by using terms such as "elaborated sentence," and "sentence functionally complete but structurally incomplete." Her terms were adopted by a number of other researchers; for example Davis²⁶ and Templin.²⁷

Because of the lasting influence of the criterion the study established, it is worthwhile to examine the work of La Brant.²⁸ With respect to identifying sentences, she wrote:

. . . it is impossible to determine what constitutes a sentence in an individual's oral or written composition, unless the sentence be perfectly punctuated by marks or inflection.²⁹

Further, with respect to words per clause, she wrote, "length of clause is not a significant measure of language development for children in grades 4 to 12, inclusive."³⁰ However, this conclusion may have stemmed in part from her definition of a clause. For example, La Brant would say that the expression, "I am trying hard but not succeeding," (which has two predicates, but only one clause in present-day practice) has two clauses--one represented by each predicate. A more modern approach might have altered her conclusions. However, she went on to define a new metric, the "subordination

²⁶Davis, op. cit.

²⁷Templin, op. cit.

²⁸La Brant, op. cit.

²⁹La Brant, op. cit., p. 482.

³⁰La Brant, op. cit., pp. 467-468.

ratio," based upon the proportion of predicates in dependent, subordinate clauses compared to the number of predicates in main clauses. This ratio or "index" she found to increase consistently with increasing age. This metric, which still appears in present-day literature, represents the first effective method of examining significant relationships within sentences.

With the great upsurge in interest in syntax brought about by linguistic science in the late 1950's and 1960's, methods based upon structural and transformational grammatical theories began to appear. Using structural approaches were the studies of Loban,³¹ Strickland,³² Hocker,³³ Riling,³⁴ and Sam and Stine.³⁵ Transformational approaches have been used by Menyuk,³⁶ Hunt,³⁷ Loban,³⁸ O'Donnell, Griffin and Norris,³⁹ and Fagan.⁴⁰

³¹Loban, op cit., 1961, 1963, 1964.

³²Strickland, op. cit.

³³Hocker, op. cit.

³⁴Mildred E. Riling, Oral and Written Language of Children in Grades 4 and 6 Compared with the Language of Their Textbooks. Report to the U.S. Office of Education, Cooperative Research Project No. 2410. Durant, Oklahoma: 1965.

³⁵Sam and Stine, op. cit.

³⁶Menyuk, op. cit.

³⁷Hunt, op. cit., 1964, 1965, 1966, 1970.

³⁸Loban, op. cit., 1967.

³⁹O'Donnell et al., op. cit.

⁴⁰Fagan, op. cit.

Strickland's study⁴¹ is an excellent example of the structural approach. Observing the frequency with which younger children use coordinate conjunctions to stretch sentences, she wrote, "Length of phonological unit appeared in this study to be unsatisfactory as a measure of the maturity of language."⁴² Her analysis of the language of children was organized into two levels using a modified "slot and filler" approach which classified items according to the "slots" they could fill, but also had classes of "moveables," "connectors," and "satellites" or subordinate constructions. The result was an extremely complex classification system based on a number of criteria. Strickland identified ten "immoveables" or slots in the language of older students that were not in the repertoires of first graders. However, she concluded that overall sentence patterns changed little. She also noted an increase in the number of adverbials used by older students, and a greater variety of positions of occurrence as well. One interesting finding of Strickland's study was her conclusion with respect to subordinate structures. She apparently found "no outstanding difference in the use of these elements from one grade level to another."⁴³ However, she had found significant variations in the use of subordinate structures at two grade levels based upon interaction of factors of

⁴¹Strickland, op. cit.

⁴²Strickland, op. cit., p. 60.

⁴³Strickland, op. cit., p. 44.

mental age, parents' education, and verbal intelligence. Strickland was one of the first investigators to note specifically that children, and especially young children, produced structures of little or no meaning, which she termed "mazes."

Loban⁴⁴ also noted the occurrence of "mazes," and used them as one of the criteria for his study in 1963. He used intonation patterns as criteria for divisions of oral language, and based his analysis of written language upon the concept of the "communication unit,"--a unit of language "which cannot be further divided without the loss of . . . essential meaning," this definition being borrowed from Watts.⁴⁵ Loban's classification system, somewhat different from Strickland's, was also very complex. He used numbered functional units such as subject, inner complement, moveables, mazes, and so forth. He further broke down the specific functions performed by units. For example, a group might be labelled "facts and unelaborated perceptions," "generalizations," or "irrelevancies."⁴⁶ Oral language was also classified according to a 1 to 3 scale on evaluative criteria such as "fluent to halting," or "mature to babyish."⁴⁷

⁴⁴Loban, op. cit., 1963.

⁴⁵A. F. Watts, The Language and Mental Development of Children. Boston: D. C. Heath and Company, 1948, pp. 65-66.

⁴⁶Loban, op. cit., 1963, p. 16.

⁴⁷Loban, op. cit., 1963, p. 17.

Loban also examined amount of subordination, usage, and vocabulary factors. The conclusions of this study which are of interest to the present investigation are summarized below:

. . . The group proficient in language employs the linking verb sentence pattern to a greater extent than does the low group. . . .

. . . The expletive type of sentence pattern is seldom used by the low group; for the high group, the use of it first increases, then decreases. . . .

. . . The outer complement pattern is used only by the most able of the high group, and the inner complement pattern (indirect object) is seldom used by either group. . . .

. . . Except for the linking verb patterns and the use of partials, the differences in structural patterns used by the two groups are negligible. This similarity in use of patterns is considered to be an important finding of this study, especially when considered in relation to the findings which immediately follow.

Although differences in structural patterns are not notable--with the exception of partials and linking verbs--very important differences do show up in the dexterity with which subjects use elements within these structures. The nominals, whether in subject or object position, and the movable elements show marked differences when high and low groups are compared. This holds true consistently for any syntactical nominal structure... Not pattern but what is done to achieve flexibility within the pattern proves to be a measure of effectiveness and control of language at this level of language development.

In the movable elements of the patterns, the high group consistently shows a greater repertoire of clauses and multiples (movables within movables).

For subject nominals, the low group depends almost exclusively on nouns and pronouns. The high group can use noun clauses, infinitives, and verbals. . . .

For nominals used as complements . . . the high group invariably exceeds the low group in the use of infinitives and clauses.

Adverb and noun clauses are used by the total group much more frequently than adjective clauses.

The adverb clause discriminates between high and low groups better than do the noun and adjective clauses.

On the idea of subordination, the high group uses this grammatical complexity to a greater extent than the random and low groups. Their precedence over the other two groups is consistent throughout all seven years of the study.

All three groups show an increasing use of subordination as chronological age increases.

Both the idea of subordination, and the transformational analysis show complexity of grammatical structure to be associated not only with chronological age, but also with proficiency in language and with socioeconomic status. . . .

The use of subordinating connectors increases with chronological age, mental ability, language ability, and socioeconomic status.⁴⁸

McCuaig, in looking at the work of Strickland, Loban, and others, and comparing their methods with those of transformationally-based studies, criticizes the structuralist approach for confusing competence and performance.⁴⁹ He also comments upon the confusion which results from the "open ended empirical" approach using mixed classes and variable terminology. He prefers the transformational approach, and a method which seeks to examine specific facets of language performance, rather than trying to make sense out of the entire linguistic universe at once. The present investigator agrees that somewhat more specific and constrained studies

⁴⁸Loban, op. cit., 1963, pp. 83-86.

⁴⁹Roger A. McCuaig, "How Not to Analyze the Syntax of Children: A Critique and a Proposal," Elementary English, 43: (May, 1970), pp. 612-618.

are now needed, and that the framework of transformational analysis and its terminology will aid such studies; but studies such as those cited above provide the starting points, the signposts, and the pushes in the right directions for such narrower studies.

The studies by Menyuk,⁵⁰ (who was working directly in and around the Massachusetts Institute of Technology) O'Donnell et. al.,⁵¹ and Fagan⁵² are clearly using transformational analysis on their data. Not so clear is the position of Hunt,⁵³ whose initial analyses are in "traditional" terms, and in terms of what he called the "T-unit," "the minimal terminable unit" of language, which would thus have to contain at least one principal clause plus modifiers and/or subordinate structures. In contrast to Menyuk,⁵⁴ who was attempting to devise transformational grammars to account for the corpuses she was collecting, Hunt used the following method (but only as supplements to his T-unit approach): he looked for indications that different phrase structure rules were being employed by his subjects (and found no such indication); he noted that older subjects tended to elaborate verb auxiliaries, but with little increase in clause length; however, he found that older, more mature writers

⁵⁰ Menyuk, op. cit.

⁵¹ O'Donnell, et al., op. cit.

⁵² Fagan, op. cit.

⁵³ Hunt, op. cit., 1965, 1966.

⁵⁴ Menyuk, op. cit.

used a much larger proportion of generalized sentence combining and embedding transformations, especially to expand nominal structures. He attributed his carefully-documented increase in T-unit and clause length to such operations.

The study by O'Donnell, Griffin, and Norris paralleled that of Hunt's 1965 study in many ways; but while Hunt had used subjects from Grades 4, 8, and 12, the O'Donnell et al. study drew subjects from Grades 1, 2, 3, 5, and 7. In his 1970 study, Hunt summarizes the findings of O'Donnell et al. as follows:

- a) At every grade interval mean T-unit length increased with age. The values for each grade are close to those reported by Hunt. . . .
- b) The number of subordinate clauses per T-unit increased at every grade interval. The figures reported by O'Donnell et al. (1967) are fairly close to those reported by Hunt, though Hunt's figure for grade 4 is slightly higher than O'Donnell's for grade 5. . . .
- c) Although the O'Donnell study did not report the number of words per clause, it did report (words per T-unit) and (clauses per T-unit) in (a) and (b) above, and from those values the number of words per clause can be calculated for each grade. The values for clause length arrived at in this way increase at each grade interval and are fairly close to those reported by Hunt.

| | G3 | G4 | G5 | G7 | G8 |
|--------------------|------|------|------|------|-------|
| T-Unit Length | | | | | |
| O'Donnell | 7.67 | ... | 9.34 | 9.77 | ... |
| Hunt | ... | 8.51 | ... | ... | 11.34 |
| Clauses per T-unit | | | | | |
| O'Donnell | 1.18 | ... | 1.27 | 1.30 | ... |
| Hunt | ... | 1.29 | ... | ... | 1.42 |
| Clause Length | | | | | |
| O'Donnell | 6.5 | ... | 7.4 | 7.7 | ... |
| Hunt | ... | 6.6 | ... | ... | 8.1 |

The preceding table from Hunt's 1970 study⁵⁵ reports the results of the Hunt 1965 study and the O'Donnell et al. study of 1967 (1968).

Hunt continues,

O'Donnell found that several of the development trends observed for writing also hold true for speech --and hold even in the earliest grades:

- a) T-unit length in speech increases at every grade interval from kindergarten to grade 1, to grade 2, to 3, to 5, to 7.
- b) The number of clauses per T-unit for speech did not increase at every interval; nonetheless the number did increase, though in a zigzag upward path. The values for kindergarten and grades 1, 2, 3, 5, and 7 were, respectively, 1.16, 1.19, 1.18, 1.21, 1.19, 1.26.
- c) The number of words per clause can be calculated from O'Donnell's published values for words per T-unit and subordinate clauses per T-unit. The values arrived at in their way increase at every successive grade. For the five grades, in order, the values are 6.1, 6.7, 7.1, 7.2, 7.5, 7.8.⁵⁶

With respect to the use of transformational analysis in O'Donnell's study, and reflecting on his own findings in 1965, Hunt writes,

The O'Donnell study also tabulated the number of sentence-combining transformations used by the students at each grade level both in speech and writing. In both modes of expression he found an increase in number at each grade interval. He concluded that the parallels between number of such transformations and T-unit length for each group were "impressive."⁵⁷

⁵⁵Hunt, op. cit., 1970, p. 9.

⁵⁶Hunt, *ibid*, pp. 8-9.

⁵⁷Hunt, *ibid*, P. 9.

Hunt concluded that the process of sentence-combining by transformational process was used more and more frequently with greater maturity. He also noted that this process would explain and underlie his earlier observations concerning increasing T-unit length.

Hunt's 1970 study was designed to explore the ways in which students in grades 4, 6, 8, 10, and 12, and both "average" and "superior" adults went about combining sentences. The instrument used, designed by O'Donnell, was an exposition on "Aluminum" written in very short, simple sentences. Part of the directions given to students read:

. . . You will notice that the sentences are short and choppy. Study the passage, and then rewrite it in a better way. You may combine sentences, change the order of words, and omit words that are repeated too many times. But try not to leave out any of the information. . . .⁵⁸

Hunt reported that his results indicated that students exhibit the same characteristics when re-writing as they do in their original writing. His earlier findings with respect to increasing clause length and T-unit length with greater maturity were confirmed. Not only did the number of embeddings steadily increase, but an interesting pattern developed: at each ascending grade level, students tended to use all the embeddings used at the previous level, and then add several more: the picture that emerged was clearly a cumulative one with respect to ability with

⁵⁸Hunt, *ibid*, pp. 11-12.

sentence-combining operations. Hunt comments that his report shows:

. . . not only do successively older writers reduce more inputs to less than a predicate, but they use a wider variety of transformations in doing so.⁵⁹

He continues,

One can scarcely escape the conclusion that as writers mature they take advantage of more and more opportunities for consolidating sentences. No doubt, diminishing returns would set in at some point along the line if carried to too great an extreme, but the evidence presented for skilled adults indicates that the point of diminishing returns is so far in advance of the point reached by schoolchildren that curriculum makers need not worry about this matter.⁶⁰

One of the perennial questions troubling critics of writing has been the matter of content. Does "what is said" account for differences in skill? Is choice of ideas--a semantic matter--the prime factor, negating findings concerning syntax? Writes Hunt,

In this experiment, all writers were constrained to say the same thing. When a writer added something not contained in the original, the whole sentence was deleted. Yet, even so, the older writers showed the same syntactic superiority they showed in free writing, suggesting that their sentences are affected by their syntactic ability, not just by what they have to say.⁶¹

One additional study, that of Loban in 1967,⁶² has confirmed most of the findings of O'Donnell et al. and Hunt. Loban reported a steady continuance of growth at the grade 10, 11, and 12 level in ability to use, what he termed, means of

⁵⁹Hunt, *ibid*, p. 53.

⁶⁰Hunt, *ibid*.

⁶¹Hunt, *ibid*, p. 54.

⁶²Loban, *op. cit.*, 1967.

"elaboration" in speech and writing, supposedly referring to the processes of embedding and sentence combining described in the other studies.

A study originally in the area of reading by Fagan⁶³ used "Cloze" procedure to determine the relative difficulty with which students in grades four, five, and six were able to comprehend portions of stories taken from basal readers, as well as sections of stories re-written to incorporate various structures of language. The present study, although it uses a slightly different instrument, extends Fagan's findings through junior and senior high school, although slightly different structures of language were used analytically. Fagan correlated both sentence and passage difficulty against the occurrence of various transformations. At the grade four level, occurrence of eleven embedding transformations correlated with sentence difficulty, while ten correlated significantly with passage difficulty. Eight embedding transformations overlapped, with "appositives," "pronoun (genitive)," and "verb + C" relating most strongly to sentences, while "WH + S obj." and "adverb expansion + S" made passages most difficult.⁶⁴ Fagan accounted for differences between effects on passage and sentence difficulty by attributing influence to context. At the grade five level, twelve transformations correlated significantly with sentence

⁶³Fagan, op. cit.

⁶⁴Fagan, *ibid*, p. 134.

difficulty, and nine with passage difficulty, with seven in common. By grade six, the respective significant correlations numbered eight and seven, with four in common. The overall picture yielded twelve embedding transformations correlating significantly with sentence difficulty, ten with passage difficulty, and nine in common. Fagan concluded, ". . . the greater (the) number of embedding transformations within a sentence, and/or passage, the more difficult that sentence and/or passage is likely to be."⁶⁵ He noted the following exceptions: "adjective," "infinitive as object," "WH + S object," and "adverb expansion + S." Correlation using the conjoining transformation did not result in any significant results except for total samples. All of the deletion transformations, however, correlated with sentence difficulty, and three with passage difficulty at the grade four level. In grade five, these figures were six and four, and by grade six, three and four. Overall, six deletion transformations correlated significantly with sentence difficulty, and four with passage difficulty. Results with simple transformations were highly mixed because of the variability of structures included under their category; however Fagan concluded that the presence of simple transformations contributed to reading ease, possibly because of the similarity of many of the operations included in this category to operations common in oral language.

⁶⁵Fagan, *ibid*, p. 139.

"Little consistency" was noted in the effects of position-shift transformations, although those common in oral language were again noted to be relatively easier.

After the above investigation, which examined the effect of the presence of various transformations, Fagan then conducted a similar examination of the effect upon sentence and passage difficulty of various transformations when their relative difficulty was taken into consideration. Using a correlation method similar to that above, he discovered that more simple than embedding transformations fell into the "most difficult 15" category. However, embedding transformations as a group were not easy; they clustered in the "medium difficulty" group. Deletion transformations were consistently hard, while conjoining ones were easy. Position-shift transformations showed little consistency. Considering groups as a whole, Fagan stated,

"It is difficult to generalize about the findings of this study on the relationship of transformations to sentence and passage difficulty. Some transformations related more to the difficulty of written language by their presence, while others related more by their difficulty. Those transformations whose difficulty correlated with sentence and passage difficulty may have had high or low difficulty indexes. Thus generalizations are best made about a transform on either of these factors separately, and for each group tested.⁶⁶

Summary

While Fagan does not generalize about classes of transformations beyond suggesting that the presence of

⁶⁶Fagan, *ibid*, p. 182.

embedding and deletion transformations contributes to passage difficulty, that simple ones contribute to easier passages, and that conjoining and position-shift transforms seem to be poor predictors, even these statements support the earlier findings of Hunt, O'Donnell et al., and Loban. It appears that the structures which contribute to greater elaboration, longer T-units, and a greater "richness" in writing--the factors which particularly Hunt and O'Donnell et al. identify as being earmarks of superior writers --are structures which children in grades 4, 5, and 6 have trouble reading, and, one might imply, using in their own written work. That deletion transformations are also somewhat difficult for them is not really a factor: indeed, absence of ability with deletion operations would merely tend to make the writer's T-units that much longer. Similarly, a writer's poor mastery of a simple transformation--for example, the passive--would simply result in his not using it. Thus the T-unit, while it may be capable of measuring a writer's ability with sentence combining operations, cannot measure all of the dimensions of a writer's ability.

If sentence length and subordination index are not good criteria, and if a structural approach cataloguing what a writer does use is too complex, and if the T-unit is an incomplete measure, and if transformational analysis leads to inconclusive results, where are we to turn? Further, what are we to teach? Clearly, the best writers, indicated particularly by Hunt's 1970 study, make use of

sentence-combining operations which are capable of operating only when the writer is able to discern relationships and contingencies within what he wishes to say which make the use of such devices possible. Must we conclude that ability as a writer is coupled with overall development, that a "readiness principle" is operating which makes any attempt at teaching a student how to be a better writer fruitless? This question will be examined in the final chapter of the present study. However, the next portion of this chapter examines several attempts at using some of the knowledge which researchers have developed already. Included is a brief outline of an approach suggested by Francis Christensen, an approach which involves methods of enriching and extending the modification system of English. Also included are two studies examining the question of whether transformational grammar can be adapted to the teaching of writing. These studies are included in this survey of literature only to illustrate that research into syntax can be applied to the teaching situation, a fact which many teachers exposed to Chomsky and his phrase markers have doubted. They are put forward also in the hope that the insights of the work of Strickland, Loban, Hunt, O'Donnell et al., and Fagan may find expression in new approaches to the teaching of writing.

B. WRITING APPLICATIONS

Bateman and Zidonis⁶⁷ conducted a two-year study

with 41 students in grades 9 and 10 in which they attempted to find answers to the following questions:

- 1) Can high school pupils learn to apply the transformational rules of a generative grammar in their writing?
- 2) Can their repertoire of grammatical structures be increased by a study of generative grammar?
- 3) To what extent will the proportion of well-formed sentences increase in pupil writing over the two-year period?
- 4) What kinds of transformational errors will occur in pupil writing, and to what extent will such errors increase or diminish over the two-year period?⁶⁸

The experimental group were taught in an intensive fashion from materials prepared by the investigators. These materials outlined the theory of transformational grammar along with incorporating exercises designed to give students practice at deriving, transforming, and combining sentences. The investigators also prepared an analytical transformational grammar which they used to classify the output of both experimental and control groups. Samples of writing were obtained from both groups during the first three months and the last three months of the two-year experiment. Student writing was judged as to structural complexity, proportion of well-formed sentences, and error changes, or misapplication of some rule of grammar. Results were as follows, all comparisons being between experimental and control groups:

⁶⁷Bateman and Zidonis, op. cit.

⁶⁸Ibid, p. 3.

before and after gains in structural complexity were significant at the .01 level, and interaction at the .05 level, but there was no significant difference between experimental and control groups. Inspection of results, however, indicated some gains for the experimental group. In respect to proportion of well-formed sentences, before and after, interaction, and comparison results were all significant at the .01 level, clearly favoring the experimental group. Error change scores were significant at the .01 level with respect to before and after, and experimental/control dimensions, but no significance was found for interaction. A further examination of the proportions of various transformations in the work of the two groups was inconclusive. These results led Bateman and Zidonis to the following conclusions, although they cautioned inferences because of the size of their experimental group:

. . . the persistently higher gains scores for the experimental class in every comparison made strengthens the contention that the study of a systematic grammar which is a theoretical model of the process of sentence production is the logical way to modify the process itself.⁶⁹

While the present investigator would question the statement above concerning a transformational grammar being a model of "sentence production," still, such results are encouraging.

⁶⁹Bateman and Zidonis, Ibid, p. 37.

Mellon,⁷⁰ however, who also questions transformational grammar as being a model of sentence production, makes a number of charges against the Bateman-Zidonis study. He questions several of the statistical techniques used, particularly the fact that investigation was not limited to items shown in earlier research to be subject to development. He wonders exactly what was done in class with the control group. More seriously, he makes the charge that Bateman and Zidonis assumed that their students would consciously apply the rules they were being taught, an assumption which much research indicates does not happen with traditional grammar rules. He also wonders why Bateman and Zidonis did not limit their treatment to exercises in sentence combining, as Bateman had suggested in an earlier paper.⁷¹ Noting the earlier research by Hunt that made such a practice reasonable,⁷² Mellon designed his own experiment along just such lines. Working with 247 grade seven students, Mellon set up an experimental group of 100, a control group of 100 receiving "normal" grammar instruction, and a placebo group of 47 who did

⁷⁰ John C. Mellon, Transformational Sentence-Combining: A Method for Enhancing the Development of Syntactic Fluency in English Composition. Final Report, Cooperative Research Project No. 5-8418, U.S. Department of Health, Education and Welfare, Office of Education, 1967.

⁷¹ D. R. Bateman, Speculations Concerning Symbolism, the Communications Core, and Language. Columbus: Center for School Experimentation, The Ohio State University, 1959.

⁷² Hunt, op. cit., 1965, 1966.

extra work in literature, but who received no instruction in grammar. Pre and post samples of writing on carefully selected and equalized topics were collected at the beginning and end of the experimental year. The experimental group worked on sentence-combining exercises and problems such as the following:

The office building towered above the tenements.
 The building was gleaming.
 The building was new.
 The building was rising high into the sky.
 The tenements were decrepit.
 The tenements were brick.
 The tenements were in the slums.
 The slums were surrounding this symbol of prosperity.
 The prosperity was supposedly universal.
 The prosperity was American.⁷³

Each indentation indicates a new level of "embedding." Students were instructed to work from the bottom upwards, combining sentences and deleting unnecessary elements, their goal being to produce:

The gleaming new office building rising high into the sky towered above the decrepit brick tenements in the slums surrounding this symbol of supposedly universal American prosperity.⁷⁴

Mellon carefully analyzed the writing from his experimental and control groups on twelve factors, and found changes significant at the .01 level on all twelve for his experimental group, and on three for his control group. Using the earlier results of Hunt as a criterion, he stated that the experimental group had experienced over one additional

⁷³Mellon, op. cit., p. 131.

⁷⁴Ibid.

year of growth in syntactic fluency over his control group. Results were the same on boy/girl and urban/suburban dimensions. One interesting development, however, was the result of a carefully-conducted writing test administered to all three groups. Six experienced markers, whose inter-marker reliability was .83, rated papers for all subjects on a five point scale using a standardized set of criteria of writing quality. Results showed the experimental and placebo groups to be indistinguishable, while the control group was rated better at the .01 level of significance. Such results appear discouraging, but one wonders whether the markers were dealing with a variety of writing none of them had ever seen before.

A theory concerning a somewhat different method of enhancing writing ability, or achieving a "more mature" style, has been put forward by Christensen.^{75,76} He speaks of a "generative rhetoric," of composition being "essentially a process of addition," of "the principle of direction of modification or direction of movement," of "cumulative sentences," of "levels of generality or levels of abstraction," and of the "texture" which writing has.⁷⁷ One of the examples which Christensen supplies of the type

⁷⁵Francis Christensen, "The Problem of Defining a Mature Style," English Journal, LVII (April, 1968), pp. 572-579.

⁷⁶Francis Christensen, Notes Toward a New Rhetoric. New York: Harper and Row, 1967.

⁷⁷Christensen, op. cit., 1967, pp. 3-8.

of writing he considers to demonstrate the above traits is the following:

1. The Beach sounds are jazzy,
2. percussion fixing the mode--(Abs)
 3. the surf cracking and booming in the distance, (Abs)
 3. a little nearer dropped bar-bells clanking, (Abs)
 3. steel gym rings/ringing, (Abs)
 - /4 flung together, (VC)
 3. Palm fronds rustling above me, (Abs)
 - /4 like steel brushes washing over a snare drum, (PP)
 3. troupes of sandals splatting and shuffling on the sandy cement, (Abs)
 - /4 their beat varying, (Abs)
 - /5 syncopation emerging and disappearing with changing paces. (Abs) ⁷⁸

The indention and numbers once again represent different "levels" of modification, the letter symbols marking the grammatical character of the additions. "Abs" stands for an absolute, "VC" for a verb cluster, "PP" a prepositional phrase; and, in addition, he suggests that the following are available: "SC" - subordinate clause, "RC" - relative clause, "NC" - noun cluster, "AC" - adjective cluster, and "A + A" - adjective series. This example of Christensen's "cumulative sentence" makes use of sentence modifiers or "free" modifiers to create a full, rich picture. Such writing, although more appropriate to descriptive and narrative writing than to exposition, is, according to Christensen, the mark of the superior or more mature writer.

⁷⁸Ibid, p. 12.

In examining the studies of Hunt⁷⁹ and Mellon,⁸⁰ Christensen notes the various measures of maturity which Hunt cites (T-unit length, clause length, subordination ratio), and then quotes Mellon twice, as commenting that Hunt has shown the ability growth factors as enabling writers to "say more" with every statement, and as saying that "it is the nominal and relative transforms whose consistently greater frequencies per T-unit characterize growth of syntactic fluency."⁸¹ Christensen claims that Mellon's experimental approach may be headed in the wrong direction, producing long noun phrases which he terms "the very hallmark of jargon."⁸² In conclusion, besides advocating shorter noun phrases, Christensen says the following in the form of hypotheses:

1. A mature style will have a relatively high frequency of free modifiers, especially in the final position. The frequency of free noun, verb, and adjective phrases and of verbid clauses will be high.
2. Such a style will have also a relatively high frequency of structures of coordination within the T-unit--what might be called intra-T-unit coordination. Inter-T-unit coordination, producing compound sentences, should be regarded as a feature of paragraph rather than sentence structure.

One study which attempted to apply the insights of

⁷⁹Hunt, op. cit., 1965.

⁸⁰Mellon, Op. cit., 1967.

⁸¹Christensen, op. cit., 1968, pp. 574-575.

⁸²Ibid, p. 575.

Christensen was that of Gajadharsingh.⁸³ He prepared materials teaching Christensen's ideas of sentence rhetoric, and applied them to grade seven students for five weeks. Pre, post, and retention tests marked by three experienced teachers indicated not only significant growth in students' ability to enhance their sentences, but retention of those same skills some time later. Further, low ability students were able to "improve" their styles every bit as well as high ability students.

General Summary

Research into what constitutes syntactic ability and writing ability was plagued in its earlier stages by superficiality, and by a great confusion of nomenclative and classification devices. Many studies noted that "better" writing was characterized by subordination and by modification through single words, phrases, and clauses. Recent studies based upon transformational approaches have made such findings more explicit, although further work remains to be done. Finally, tentative attempts have been made to apply some of the knowledge which research has provided about syntactic ability with some encouraging results.

⁸³Joel Lancelot Gajadharsingh, "A Study of the Effects of Instruction in the Rhetoric of the Sentence on the Written Composition of Junior High School Students." (Unpublished Doctoral dissertation, University of Alberta, 1970).

5. THE EVALUATION OF WRITING

Because this study incorporates a writing sample, and because one particular system of evaluation has been adopted, it seems necessary to justify that choice, and to review the rather scanty literature supporting it, and establishing it as a valid and reliable criterion.

The most common approach to evaluation of student writing is the score sheet, a device which breaks the total number of marks available on a writing assignment into smaller proportions. Usually such "sheets" are split into two sections, one for mechanics or style, and the other for content. An example of such a device is that used by Buxton.¹ It is possible for markers using such a method and undergoing a training period to achieve very consistent results. (For example, Buxton reported an inter-marker reliability on pretest themes of .91 and of .88 on post-test themes.)² However, out of 300 marks to be given each theme, the Buxton sheet gives only 10 marks for unity within paragraphs, 10 for transition within paragraphs, 10 for general coherence, 20 for variety in sentence structures,

¹Earl W. Buxton, "An Experiment to Test the Effects of Writing Frequency and Guided Practice upon Students' Skill in Written Expression" (Unpublished Doctoral dissertation, Stanford University, 1958) in Richard Braddock, Richard Lloyd-Jones, and Lowell Schoer, Research in Written Composition. Champaign, Ill.: National Council of Teachers of English, 1963, p. 64.

²Ibid, p. 66.

and 10 for general fluency, a total of 60 marks, or 20 percent of the marks available, for factors which much research seems to indicate are the key indicators of "greater writing maturity." However, the biggest problem of the score sheet or "analytic" approach seems to be its inflexibility, its legislative quality, in that it seems to set up absolute standards of "correctness," a practice reminiscent of prescriptive grammar. Further, such an approach cannot differentially reward an unusual gift in one aspect of writing--for example, superb organizational skill--because of the limitation of the sheet. For these reasons, a marking sheet approach was not chosen for the present study.

A second approach is the "writing scale," a carefully graded set of papers which represent a number of levels (usually 10) of relative ability in written composition. One study using such a scale was that of Smith.³ The principal problem of the scale method is that it is extremely difficult to make comparisons between any two compositions written by different people. A rater using such a method tends to pick out a number of criteria in arbitrary fashion, and to ignore other factors. Consequently, the artificiality of the "scale" method caused the present investigator to reject it.

One system often used in classrooms as a timesaver

³Dora V. Smith, Class Size in High School English: Methods and Results. Minneapolis: University of Minnesota Press, 1931.

and as a teaching device is that of concentration upon only one or two features of an assignment (for example, carefully marking pronoun usage after a lesson on pronouns). Such a specialized approach seemed unworkable for the present study.

The method finally selected was the "general impression" or "wholistic" method, whereby markers are instructed to look for general factors such as "ability to use language as an effective tool of communication," or "overall writing fluency." Markers are also instructed to work quickly, spending no more than two minutes on a paper. The final mark for any one paper is the mean (or total) of the marks of the entire group of raters. A report by Wiseman⁴ of extensive use of the general impression method reports inter-marker reliabilities among four markers in the lower .90's. The task of Wiseman's markers, however, was to assess whether a student could profit from a grammar school education. Thus papers were divided on a simple yes/no basis, and only a few papers at the dividing point would cause serious disagreement. Such a bipart system would boost reliabilities, because a large percentage of the papers would cause no disagreement at all. Wiseman argues in his article that some disagreement among markers is a good thing, because the composite mark would then reflect a number of points of view. While some disagreement might be attributable

⁴Stephen Wiseman, "The Marking of English Composition in Grammar School Selection," British Journal of Educational Psychology, XIX (November, 1949), pp. 200-209.

to inconsistent marking, the present investigator believes that experienced markers would reduce this possibility. A study in which the same markers would mark and then remark the same set of papers at a later time might establish the validity of this charge of individual inconsistency. In a study examining four different marking methods, Cast⁵ found the score sheet or analytic method and the general impression method to be the most reliable. She praised the analytic method precisely for the way in which its criteria can be standardized, one of the reasons for rejecting that approach in the present study. Finlayson⁶ criticized Wiseman's practice of pooling marks; he claimed that reliabilities would be accurate only if a test-retest or parallel test approach was used. In an experiment using parallel tests and six markers, he obtained individual test-retest marker reliability figures ranging from .601 to .798, with a mean of .691. His test-retest reliability of pooled marks resulted in a reliability of .863. Nisbet,⁷ returning to Wiseman's practice of pooling marks obtained reliability figures of .96 for four markers on a test-retest situation.

⁵B. M. D. Cast, "The Efficiency of Different Methods of Marking English Composition," British Journal of Educational Psychology, IX (November, 1939), pp. 257-269, and X (February, 1940), pp. 49-60.

⁶D. S. Finlayson, "The Reliability of the Marking of Essays," British Journal of Educational Psychology, XXI (1951), pp. 216-234.

⁷J. D. Nisbet, "English Composition in the Secondary School Selection," British Journal of Educational Psychology, XXXV (1955), pp. 51-54.

In 1956, Wiseman reported using a test-retest design, and obtained reliabilities with four markers of .896. He also reported a method of maintaining inter-marker and intra-marker reliability: three months after a sample was marked, a 1/10th random sample of papers was remarked by all markers. Any marker not achieving a test-retest reliability of .70 was not rehired. One large study in the United States was conducted by Diederich;⁸ 53 markers from all endeavours were asked to grade 300 essays written by undergraduates into nine piles, with no less than six papers to a pile. Variations on papers were extreme. The mean correlation for all markers was 0.31, and 0.41 for teachers of English. The instructions given were very general.

Summary

For the purposes of this study, which seeks to relate syntactic ability with overall ability with language, the general impression approach--with no specific instruction to markers to look for sentence structure variety--seems a valid choice of method. It has been fairly widely used, and has proven to be capable of reliability. Its validity, however, is increased if a test-retest or parallel tests approach is used. The primary reason for choosing the general impression method, however, lies in its emphasis on the grosser structures of language, rather than on the details of mechanics and word choice which characterize the analytic method.

⁸P. G. Diederich, Factors in the Judgment of Writing Ability. E. T. S. Research Bulletin 61-65. Princeton, N.J.: E. T. S., 1961.

CHAPTER III

THE GRAMMAR OF THE STUDY

The analyses performed in the principal part of this study are based upon tabulations of the responses of students in grades seven, nine, and eleven to a "cloze and delete" test. The test was written to incorporate three instances each of occurrence of certain structures of language. In the past, these structures were given names based upon their function and the words they contained: for example, an adjective clause might be described as a group of words containing a subjective element and a predication, and functioning to "modify" (describe, add information about, limit the meaning of, and so forth) some nominal element. Thus the sentence, "The boy who is standing on the corner is my friend " would be said to contain a restrictive relative adjective clause, ". . . who is standing on the corner . . . ," which limits the meaning of "boy" to the one standing on the corner, and tells just what that boy is doing. A similar function can be seen for ". . . with a red cap on . . . " in "The man with a red cap on is a hunter." Never in the early studies of syntactic capabilities and development was it possible to use a term capable of including all such elements; they were often enumerated as being the marks of "better" or "more mature" writers. However, modern generative-transformational grammar has provided us with such a tool.

Transformations, in an Aspects or post-Aspects grammar, are either obligatory or optional structural changes which are purely interpretative in nature. They apply to products of the base component having or meeting specified structural descriptions.

The recursive element of such a grammar appears in the base component by means of the inclusion of optional "S" symbols (representing embedded sentence structures). Thus the sentence, "The man who is in the picture is my uncle" begins as "The man is my uncle" with an optional "S" appearing, attached to the "NP" node dominating "the" and "man." This embedded sentence would be something like, "The man is in the picture," which would become "who is in the picture" by the relativization transformation.

Cyclic application of transformational rules, which apply to the most deeply embedded sentence first, and then progressively up the phrase-structure tree, would then embed or attach the constituent sentence into the matrix sentence, yielding the final form, or surface structure.

Transformations, then, may be seen to be major operations which give final shape to, or interpret the output of the base component. They may give shape to elements such as negation; they may delete items; they may conjoin items; they may shift the positions of items.

While some generalizations concerning the relative difficulty of certain of these major operations may be made in terms of the results of this study, it must be realized

that great variations in relative difficulty of transformations within the major categories must be anticipated. For example, while a transformational grammar makes the insertion of a prenominal adjective a very complex thing indeed, involving processes of relativization, embedding, deletion, and front-shifting, in practice subjects who plug adjectives in front of nouns frequently every day of their lives will surely find such operation very easy. As another example, the processes of passivization and negation are both classified as "simple" transformations, but one certainly seems more complex than the other. As a result, one would expect overlap between certain of the major classifications, and great variations of relative difficulty of structures within classes.

The grammar adopted for this study is derived from Chomsky's basic model in Aspects of the Theory of Syntax.¹ Modifications include the suggestions for incorporation of elements such as Q, NEG, and IMP within the base structure made by Katz and Postal in An Integrated Theory of Linguistic Descriptions.² Further modifications have come from reference to Jacobs and Rosenbaum's English Transformational Grammar,³ and Rosenbaum's The Grammar of English Predicate Complement

¹Chomsky, op. cit.

²Jerrold Katz and Paul M. Postal. Cambridge, Mass.: The M.I.T. Press, Research Monograph No. 26, 1964.

³Roderick A. Jacobs and Peter S. Rosenbaum. Waltham, Mass.: Blaisdell Publishing Company, 1968.

Constructions.⁴ Finally, changes within the base structure and lexicon are directly attributable to class work and discussions with Dr. Gary Prideaux, Department of Linguistics, University of Alberta.

GLOSSARY OF GRAMMAR

| | | |
|-----------------|-------|--|
| S | = | Sentence |
| NP | = | Noun Phrase |
| VP | = | Verb Phrase |
| Q | = | Question |
| NEG | = | Negative |
| AUX | = | Auxiliary |
| PP | = | Prepositional Phrase |
| ADV | = | Adverb |
| P | = | Preposition |
| D | = | Determiner |
| T | = | Tense |
| M | = | Modal |
| PL | = | Plural |
| VB | = | Verbal |
| PROG | = | Progressive |
| PAST | = | Past Tense |
| WH | = | Relative Pronoun "Dummy" |
| HUM | = | Human |
| AN | = | Animate |
| CNT | = | Count |
| ABS | = | Abstract |
| STAT | = | Stative |
| ML | = | Male |
| Square Brackets | [] | = Features or Feature Matrices |
| Parentheses | () | = Optional Elements |
| Braces | { } | = Alternate Elements |
| | /_(x) | = "In the environment of following an optional x." |
| | → | = "May be rewritten as . . ." |
| | ^ | = Concatenation |

⁴Peter S. Rosenbaum. Cambridge, Mass.: The M.I.T. Press, Research Monograph No. 47, 1967.

POST-ASPECTS GRAMMARTHE BASE COMPONENT

Branching Rules

1. S → (Q) (NEG) NP[^]AUX[^]VP
2. VP → V (NP) (PP) (ADV)
3. PP → P[^]NP
4. NP → $\left\{ \begin{array}{l} \text{NP}^{\wedge}\text{S} \\ (\text{D}) \text{N}(\text{S}) \end{array} \right\}$
5. AUX → T (M)
6. X → [+X] where X = {N, T, V, M, P, D, Neg, Q}

(Actually, Rule 6 is merely a general convention, and not a rule at all.)

Grammatical Feature Rules

7. [+N] → [\pm PL]
8. [+V] → [\pm VB, \pm PROG]
9. [+T] → [\pm PAST]
10. [+D] → [\pm WH]

LEXICON

Lexical Redundancy Rules (To be applied to
Lexical entries before
Lexical Insertion is
made)

1. [+HUM] → [+AN]
2. [-AN] → [-HUM]
3. [+CNT] → [-ABS]
4. [-COM] → [-CNT, -ABS]
5. [-CNT] → [-PL]
6. [+STAT] → [-PROG]

Examples of PossibleLexical Entries:

| | | |
|--------------|-----------------------------|----------|
| <u>boy</u> | [+N, +COM, +CNT, +HUM, +ML] | / (D) __ |
| <u>girl</u> | [+N, +COM, +CNT, +HUM, -ML] | / (D) __ |
| <u>dog</u> | [+N, +COM, +CNT, -HUM, +AN] | / (D) __ |
| <u>tree</u> | [+N, +COM, +CNT, -AN] | / (D) __ |
| <u>book</u> | [+N, +COM, +CNT, -AN] | / (D) __ |
| <u>pen</u> | [+N, +COM, +CNT, -AN] | / (D) __ |
| <u>chair</u> | [+N, +COM, +CNT, -AN] | / (D) __ |
| <u>milk</u> | [+N, +COM, -CNT, -ABS, -AN] | / (D) __ |
| <u>wheat</u> | [+N, +COM, -CNT, -ABS, -AN] | / (D) __ |

| | | | |
|--------------|------------------------------------|---------------|----------------------|
| <u>John</u> | [+N, -COM, +HUM, + ML] | / | — |
| <u>Mary</u> | [+N, -COM, +HUM, -ML] | / | — |
| <u>Fido</u> | [+N, -COM, -HUM, +AN, +ML] | / | — |
| <u>fact</u> | [+N, +COM, +CNT, -AN, +ABS, +FACT] | / (D) | — |
| <u>story</u> | [+N, -COM, +CNT, -AN, +ABS,] | / (D) | — |
| <u>the</u> | [+D, +DEF, -WH] | } | / __ [+N, -PL] |
| <u>a</u> | [+D, -DEF, -WH] | | |
| <u>some</u> | [+D, -DEF, -WH] | | |
| <u>watch</u> | [+V, +VB, -STAT] | / __ NP (PP), | [+AN] _ [-ABS] |
| <u>see</u> | [+V, +VB, -STAT] | / __ NP (PP), | [+AN] _ [-ABS] |
| <u>walk</u> | [+V, +VB, -STAT] | / __ (PP), | [+AN] _ |
| <u>smile</u> | [+V, +VB, -STAT] | / __ (PP), | [+HUM] _ |
| <u>run</u> | [+V, +VB, -STAT] | / __ (PP), | [+AN] _ |
| <u>own</u> | [+V, +VB, +STAT] | / __ NP | [+HUM] _ [-ABS] |
| <u>give</u> | [+V, +VB, -STAT] | / __ NP (PP), | [+AN] _ [-ABS] [+AN] |
| <u>buy</u> | [+V, +VB, -STAT] | / __ NP (PP), | [+AN] _ [-ABS] [+AN] |
| <u>can</u> | [+M] | | |
| <u>will</u> | [+M] | | |

| | | |
|---------------|--|--------------------------|
| <u>must</u> | [+M] | |
| <u>happy</u> | [+V, -VB, -STAT] | /__ (PP), [+AN]__ |
| <u>ill</u> | [+V, -VB, +STAT] | /__ (PP), [+AN]__ |
| <u>tall</u> | [+V, -VB, +STAT] | /__ [-ABS]__ |
| <u>eager</u> | [+V, -VB, -STAT] | /__ (PP), [+AN]__ |
| <u>easy</u> | [+V, -VB, +STAT] | /__ (PP), [+ABS, -CNT]__ |
| <u>pretty</u> | [+V, -VB, +STAT] | /__ [-ABS]__ |
| <u>sick</u> | [+V, -VB, -STAT] | /__ (PP), [+AN]__ |
| <u>in</u> | [+P], <u>on</u> [+P], <u>under</u> [+P], <u>to</u> [+P], <u>from</u> [+P], | |
| <u>below</u> | [+P] | |
| <u>not</u> | [+NEG] | |
| <u>∅</u> | [+N, +PRO, {+HUM, -AN, +AN, +LOC, +TM, +COM, -CNT}] | |

The Transformational Component

The following list of transformations represents the structures incorporated in the "cloze" test. A more complete list of the transformations follows. This list does not claim to be a comprehensive one; it is given as representative of the techniques of normal transformational technique, and is subject to challenge and refinement.

Transformations of Cloze Test

1. Relative Clause (Subject)
2. Relative Clause (Object)
3. Pre-Nominal Adjective
4. Gerundive
5. "with" Phrase
6. Genitive
7. Appositive
8. It-That (Subject) Complement
9. It-That (Object) Complement
10. For-To (Subject) Complement
11. For-To (Object) Complement
12. Possessive-"Ing" (Subject) Complement
13. Possessive-"Ing" (Object) Complement
14. Complement Deletion
15. Factive Insert
16. WH Complementiser Insert
17. If-Whether Insertion
18. "That plus S as Object Deletion (Quotation)
19. WH Deletion
20. WH-"Be" Deletion
21. Common Elements Deletion
22. Adverbial Replacement Deletion
23. Adverbial Replacement
24. Adverbial Expansion of Manner plus C
25. Conjunction
26. Passive

27. Here-There Inversion
28. Dative Movement
29. Particle Movement
30. Noun Phrase-Verb Inversion
31. Adverbial Position Shift
32. Adverbial Replacement Shift
33. Negative
34. Yes/No Question
35. WH Question
36. Imperative
37. Pronominalization
38. Reflexive

Transformations of the Grammar

The transformations of the grammar of this study have been adapted with some modifications from four principal sources:

1. Noam Chomsky, Syntactic Structures. The Hague: Mouten & Co. 1965.
2. Donald Bateman and Frank Zidonis, The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Graders. Champaign, Illinois: NCTE, 1966.
3. Roderick A. Jacobs and Peter S. Rosenbaum, English Transformational Grammar. Waltham, Massachusetts: Blaisdell Publishing Company, 1968.
4. Gary Prideaux, (Unpublished sample grammars, Linguistics 404, University of Alberta, 1970).

The basis for selection of transformations for this study has been primarily a question of frequency of common occurrence of structures which the transformations included in this compilation will account for. The definitive transformational-generative grammar of English is yet to be written, and probably never will be, and this very small grammar is acknowledged to be fragmentary at best. It has served, however, to generate the structures of the "Cloze" test which constitutes the main instrument of this study. The transformations which appear here are typical: For each a structural description is provided. They are applied in the following manner: the phrase marker (or "tree diagram") of a sentence is scanned to see if the structural description is met. If it is met, obligatory transformations must be applied, while optional transformations may be applied if the resulting structure is what is desired. Note that transformations can rearrange or delete elements, or add morphemes, but they cannot create basic phrase marker nodes.

Transformations

Note: The T rules apply cyclically, beginning with the deepest embedded S.

Also note: Transformations of the "Cloze" test are indicated by the word (Included).

T 1. Passive (Optional) (INCLUDED)

SD: NP X V NP
 SC: 1 2 3 4 →
 4 2 be+EN+ 3 by+ 1

COND: 1≠4

Example: Sam was bitten by the dog.

T2 BE Insertion

SD: AUX [+V
-VB]
 SC: 1 [α PROG]
 2 →
 1 [+V
+VB
α PROG
be] ^2

Example: John was handsome.

T3 PROGRESSIVE

SD: AUX [+V
+VB]
 SC: 1 [+PROG]
 2 →
 1 [+V
+VB
be] ^ 2 [+AF
ing]

Example: The horse was winning the race.

T4 Aux Incorporation

SD: T $\left[\begin{array}{c} +V \\ \{ \text{have} \} \\ \underline{\text{be}} \end{array} \right]$

SC: 1 2 \longrightarrow
 $1^{\wedge}2$ null

Example: Had Harry found the jewels?

T5 AGREEMENT

SD: $\left[\begin{array}{c} +N \\ \alpha PL \end{array} \right]$ $[+T]$

SC: 1 2 \longrightarrow
 1 $\left[\begin{array}{c} 2 \\ \alpha PL \end{array} \right]$

Example: The boy has thrown the bomb.

The boys have thrown the bomb.

T6 NEG PLACEMENT (INCLUDED)

SD: NEG NP AUX [X]

SC: 1 2 3 VP 4 VP \longrightarrow

null 2 3 1 \wedge 4

Example: Charlie had not gone.

T7 NEG INCORPORATION (OPTIONAL)

SD: [T X] NEG
 AUX AUX

SC: 1 2 3 →
 1 2 ^ 3 null

Example: Charlie had not gone.

T8 QUESTION (INCLUDED)

SD: Q NP AUX

SC: 1 2 3 →
 3 2 null

Example: Can Prunella play?

T9 RELATIVIZATION (INCLUDED)

SD: NP [X [[+D] [+N] Y] Z]
 S NP NP S

SC: 1 2 3 4 5 6 7 →
 1 2 3 [4] [5] null 7
 [+WH] [+PRO]

Cond: 1=3

Example: The boy who is crying is ill.

T10 WH ATTRACTION (INCLUDED)

SD: [X NP Y]
S S

SC: 1 2 3 →
2^1 null 3

Cond: NP dominates +D +Z
2 +WH

Example: Who did Sam go to the store with?

T11 ADJECTIVE SHIFT (Pre-nominal Adjective) (INCLUDED)

SD: [X N V]
NP NP

SC: 1 2 3 →
1 3^2 null

Example: A handsome lad filled the door. (by T. Rel. CL, T-WH Deletion, and T-Adjective Shift).

Other forms based upon Relative Clause:

Gerundive: Example: A smiling girl met John. (By T. Relative Clause, T-WH Deletion, and T Adjective Shift). (INCLUDED)

"WITH" PHRASE Example: A book with an index is needed. (By T. Rel. Clause, and a low order rule changing WH-PRO+have to "with"). (INCLUDED)

GENETIVE: Example: The Horse's mouth/the mouth of the horse. (By T. REL. CLAUSE, T WH ATTRACTION, and a low order rule changing "which the horse has" to either of the above forms.) (INCLUDED)

T CONJUNCTION (INCLUDED)

SD: [X NP VP Y] [W NP VP Z]
 S_1 S_2

SC: 1 2 3 4 5 6 7 8 →
 1 2 3 4 {and
 but
 etc.} 5 6 7 8

Example: Joe went to the store, and he saw Martha there.

T NON-RESTRICTIVE CLAUSE (INCLUDED)

SD: [NP VP] [NP VP]
 S_1 S_2

SC: 1 2 3 4 →

[S_1 S_2 [3=4] null 2]

Examples: The boy who is deaf sang.

APPOSITIVE: Example: The boy, his son, is a golfer. (By T. CONJ., T NON-RESTRICTIVE CL.; and T WH-BE DELETION). (INCLUDED)

COMPLEMENT STRUCTURES:

TRANSFORMATIONS (After Rosenbaum, 1967).

COMPLEMENT PLACEMENT TRANSFORMATIONS:

T COMP. PLACEMENT:

A. SD: X N [NP+Y] Z
S

SC: 1 2 3 4 →
1 2 that+3 4

Example: They doubt it (that) you will go.

B. SD: X N NP+VP Z

SC: 1 2 3 4 →
1 2 { for-to } +3
{ poss-ing } 4

Example: I would hate it (for) John (to) lose the race.

C. SD: X V NP NP+VP Z

SC: 1 2 3 4 5 →
1 2 3 { for-to } +4
{ poss-ing } 5

Example: Everyone prefers (it) (for) you (to) drive slowly.

D. SD: X { for-to } NP { V } Y
{ poss-ing } { have }
be

SC: 1 2 3 4 →
1 2 3 2+4

Example: I believe (it for) John's eating to have been messy.

T SUBJECT-OBJECT INVERSION (INCLUDED)

| | | | | | | | |
|-----|---|----|-----|---|----|------|-----|
| SD: | X | NP | AUX | V | NP | S | Y |
| SC: | 1 | 2 | 3 | 4 | 5 | 6 | 7 → |
| | 1 | 5 | 3 | 4 | 6 | to+2 | 7 |

Examples: The meat tastes salty to me. (From: I taste the meat. The meat be salty.)

T EXTRAPOSITION (OPTIONAL) (INCLUDED)

| | | | | |
|-----|---|-------------------------------|------|-----|
| SD: | X | N | S | Y |
| | | { +PRO IT-THAT FOR-TO } | | |
| SC: | 1 | 2 | 3 | 4 → |
| | 1 | 2 | null | 4+3 |

Example: John is believed by me to have convinced Bill.

T COMPLEMENTIZER DELETION (INCLUDED)

| | | | | | | |
|-----|---|--------------|-------------------------------|---|----|-----|
| SD: | X | { V ADJ } | { a) [N +PRO] b) (NP) | { <u>for-to</u> <u>poss-ing</u> <u>that</u> } | NP | Y |
| SC: | 1 | 2 | 3 | 4 | 5 | 6 → |
| | 1 | 2 | 3 | null | 5 | 6 |

Examples: (It) That you came early surprised me.
John happened (for) to find gold.

T PRONOUN REPLACEMENT

| | | | | | | | |
|-----|---|----------|------|--|--------|---|------|
| SD: | X | N PRO | (AUX | $\left. \begin{array}{l} V \\ \text{be ADJ} \\ [+PROG] \end{array} \right\}$ | (MAN)) | $\left. \begin{array}{l} \text{for-to} \\ \text{poss-ing} \end{array} \right\}$ | Y |
| SC: | 1 | 2 | 3 | 4 | 5 | 6 | 7 → |
| | 1 | 7 | 3 | 4 | 5 | 6 | null |

Example: John's falling to his death was strange.

T PRONOUN DELETION

| | | | | | | |
|-----|---|---|--|---|-----|--|
| SD: | X | $\left[\begin{array}{l} N \\ +PRO \end{array} \right]$ | $\left. \begin{array}{l} \text{a) null} \\ \text{b) ADV} \end{array} \right\}$ | S | Y | Note: a) has exceptions b) is usually opt |
| SC: | 1 | 2 | 3 | 4 | 5 → | |
| | 1 | null | 3 | 4 | 5 | |

Example: That John left early angered us.

T PROPOSITION DELETION

| | | | | |
|-----|---|------|---|-----|
| SD: | X | PREP | $\left[\begin{array}{l} N \\ +PRO \end{array} \right]$ | Y |
| SC: | 1 | 2 | 3 | 4 → |
| | 1 | null | 3 | 4 |

Example: I am aware that John left.

T COMPLEMENTIZER DELETION (OBLIGATORY AFTER T PRONOUN REPLACEMENT) (INCLUDED)

| | | | | |
|-----|---|--|----|---|
| SD: | X | $\left\{ \begin{array}{l} \text{it-that} \\ \text{for-to} \\ \text{poss-ing} \end{array} \right\}$ | VP | Y |
|-----|---|--|----|---|

T-THAT plus S as OBJECT (QUOTE) (INCLUDED)

| | | | | | | |
|-----|----|---|----|-------------|-----|---|
| SD: | NP | V | NP | <u>that</u> | S | |
| SC: | 1 | 2 | 3 | 4 | 5 | → |
| | 1 | 2 | 3 | null, | "5" | |

Plus morphophonemic rules to insert punctuation and give pronouns correct shape.

Example: She told them, "Your sister is hurt."

T WH (or RELATIVE PRONOUN) DELETION (INCLUDED)

| | | | | | | |
|-----|----|---|----|--------|-----|-----|
| SD: | NP | V | NP | WH | PRO | X |
| SC: | 1 | 2 | 3 | 4 | 5 | 6 → |
| | 1 | 2 | 3 | —null— | | 6 |

Example: He has a book he wants to show you. (by T-REL. CL, T-WH ATTRACTION).

T-COMMON ELEMENTS DELETION (INCLUDED)

Any repeated element may be deleted:

Examples: His room seemed lonely and (his room seemed) damp. Joe didn't go home because he was afraid to (go home). Miriam wanted (Miriam) to leave home.

T-ADVERBIAL REPLACEMENT DELETION (INCLUDED)

Example: Sam went to the store. After that . . .
(That replacing or deleting an element in common with the previous sentence).

{ here } 2, 1 null 3
 { there }

Example: There is a thief among us.

T DATIVE MOVEMENT (INCLUDED)

| | | | | | | |
|-----|----|---|----|--------------------|----|---|
| SD: | NP | V | NP | { P for to } | NP | |
| SC: | 1 | 2 | 3 | 4 | 5 | → |
| | 1 | 2 | 5 | null | 3 | |

Examples: The dispatcher handed the engineer a message.

T PARTICLE MOVEMENT (OPTIONAL) (INCLUDED)

| | | | | | |
|-----|----|---|------|-------|---|
| SD: | NP | V | PART | NP | |
| SC: | 1 | 2 | 3 | 4 | → |
| | 1 | 2 | | 4 + 3 | |

Example: Sam held the ball up.

T NP-V INVERSION (QUOTE) (INCLUDED)

| | | | | |
|-----|----|---|---|---|
| SD: | NP | V | S | |
| SC: | 1 | 2 | 3 | → |
| | 3 | 2 | 1 | |

Example: "That's right," said Randy.

T ADVERBIAL POSITION SHIFT (INCLUDED)

| | | | |
|-----|-----|----|------|
| SD: | NP | VP | ADV |
| SC: | 1 | 2 | 3 → |
| | 3+1 | 2 | null |

Example: After a crash, they make pilots fly again.

ADVERBIAL REPLACEMENT SHIFT (INCLUDED)

Example: (You close the door (you leave at Wh TIME))
 goes to
 When you leave, close the door. (by T. REL
 CLAUSE, T IMPERATIVE, T WH ATTRACTION AND
 T EXTRAPOSITION)

T-IMPERATIVE: (INCLUDED)

T IMPERATIVE

| | | | | | |
|-----|------|-------|-----|---------------|-----|
| SD: | INP | [you] | AUX | [+V -STAT] | X |
| SC: | 1 | NP 2 | 3 | 4 | 5 → |
| | null | null | 3 | 4 | 5 |

Example: Come here. Close the door.

T REFLEXIVE (INCLUDED)

| | | | |
|-----|--------------------------|---|--------------------------|
| SD: | [N +M +F +PL] | V | [N +M +F +PL] |
|-----|--------------------------|---|--------------------------|

CHAPTER IV

THE TEST INSTRUMENT

A. GENERAL DESIGN

The test instrument of this study is based upon the theory of "Cloze procedure." This approach was chosen for the following reasons:

1. It seems to examine subjects' overall language fluency.
2. While only one word might be deleted from a sentence, the correct completion calls upon sensitivity to structure and context. That is, dependence is not solely upon knowledge of single words.
3. It avoids the difficulty of subjects' not being able to interpret questions.
4. It is relatively easy to design and score a "Cloze" test.

These last advantages were somewhat reduced by the design of the test of this particular study, and by the marking method used.

The thirty-eight transformational structures chosen to be test items (and outlined in the previous chapter) were each incorporated three times in the passages written by the examiner for this study. Then, words which formed vital parts of each such structure were deleted. For example, from the

sentence, "Sam was hit by the ball," either "was" or "by" might be deleted. Note that these two words are introduced by the passive transformation itself. A further example-- this one incorporating an "it-that" complement structure, would be, "_____ Joe was an idiot was obvious." In most cases, the elements deleted were "function" words-- words of relatively high redundancy in English. Research previously reviewed indicates that removal of function words has little effect upon comprehension.¹ The deletion items consisted of "full" sentences including redundancies and unnecessary elements. An example might be, "Bill was full of woe and Bill was full of misery." A "common elements deletion" applied to the underlined portion of this sentence would reduce the entire sentence to, "Bill was full of woe and misery." A similar operation applied to, "I saw the house which was on fire," an example of a "WH-be" deletion.

One important difference between the "cloze" and "delete" items caused the investigator some concern about the equivalence of these test items. On a "cloze" item, the subject merely has context and his own language ability to help him; he must pull a correct (or acceptable) response virtually "out of the air." All he knows is that he must insert a single word in the blank facing him. The "delete" item, however, supplies much more: it is akin to a

¹Bradley, op. cit.

multiple-choice question, in that everything is supplied. The subject knows that a word (or several words) from a group of four or five must go. He must only decide, "Which?" Might the entire class of deletion items be too easy?

Despite the doubts outlined above, the investigator could see no practical way of including deletion items other than in the fashion described above. If, in fact, people do think in terms of fuller texts which are reduced by deletion processes into surface form, then the above test would seem to capture this process adequately.

The four test passages written by the investigator were on topics of general interest-- humorous narrative, reminiscence, character study, and expository argumentative, and were written so as to include quotations and so forth, so that all the test syntactic structures could be incorporated. Four different topics were chosen, as research seems to indicate that mode of discourse may affect sentence structure,² and specific topics can also have an effect on the quality of writing (and, supposedly, response to writing).^{3,4} Furthermore, as the test was of considerable length (and researchers with

²J. C. Sagers, "Form of Discourse and Sentence Structure," English Elementary Review, X (March, 1933), pp. 51-54.

³Stephen Wiseman and Jack Wrigley, "Essay Reliability: The Effect of Choice of Essay Title," Educational and Psychological Measurement, XVIII (Spring, 1958), pp. 129 - 138.

⁴Gerald L. Kincaid, "Some Factors Affecting Variations in the Quality of Students' Writing." (Unpublished Ed. D. Dissertation, Michigan State University, 1953).

"cloze" report informally a frustration/boredom factor), it was felt that four quite different items might average out varying topic response effects, and help to eliminate potential boredom. A total of six stories or selections were originally written, the two most lacking in freshness and reading interest being discarded. Then some revisions were made so that the test items could be incorporated. A careful reading of all four passages revealed very few words which would be classified as "hard" or "obscure." Several of these were eliminated, and it was decided that the few remaining could be explained at the time of administration. Further reflection resulted in a decision to tell the subjects to ask the examiner for explanations of any word or term appearing in the passages with which they were not familiar, so as to eliminate as much as possible problems in simple comprehension.

B. PILOT STUDY

The first draft of the complete "cloze and delete" test was submitted in duplicated form to fifteen subjects ranging in achievement level from grade six to Professor. Ten of the subjects were adults skilled in the use of English. The rationale for selecting a range of subjects was to determine whether the reading level of the test material was appropriate, and whether a variety of response would be obtained. Further it was postulated that skilled adults should be capable of perfectly acceptable responses on all test items. Results confirmed that younger subjects had more

trouble with test items, although one subject in grade nine did somewhat better than another less-skilled subject in grade ten. Three test items received inconsistent responses across the entire test group. These items were re-written, and the entire test was submitted to an additional six skilled adults. Final results showed only two errors on the entire six tests, both of these appearing to be the result of careless or quick reading.

Several procedural and mechanical changes were also made as a result of pilot testing. Several subjects missed "delete" items completely. Therefore the word DELETE in capital letters was placed after every such test item. The dots indicating blanks were confused with terminal punctuation, so a slash(/) was placed at the end of each sentence. Several subjects insisted upon putting in their own punctuation, thus making incorrect responses semantically acceptable. To eliminate this practice, the test instructions were changed to forbid addition of punctuation, and the examiner also reminded subjects of this. Several subjects felt that one of the selections, "Peace in Our Time?", was somewhat too "heavy" for an opening passage; consequently, it was placed last, and "Charlie's Chore," a humorous narrative, was placed first. Instructions were also somewhat expanded, and three examples of each type of test item were included instead of one. The time required to complete all items ranged from just under one hour (grade six) to 17 minutes (a graduate student), so

the proposed administration time of 75 minutes was judged adequate to ensure every subject trying every item. The final form of the test was carefully typed and multilithed, its format and type face being judged clear and easy to read.

C. ADMINISTRATION

The final printed "cloze and delete" tests were administered to classes in their regular English language arts periods in their own classrooms. Subjects were seated in such a fashion that looking at each other's papers would be difficult, and they were asked to co-operate by not looking at their neighbor's papers. They were reminded that they were not being graded on results, and that "peeking" would only destroy the value of the research being conducted. Once the test was underway, it was noticed that better students went much faster than poorer ones, so that they were seldom working on the same page. Consequently, it was felt that "cheating" was reduced to a minimum.

Instructions given prior to administration included an admonition to read each sentence in its entirety before attempting to fill in the blanks, and to re-read each sentence after completion. Further instructions and examples appeared on the cover sheet of the test (see Appendix B), and every subject seemed clear as to what was expected. Nevertheless, after five minutes or so of working time, the investigator checked each student's answers, and found a total of three

subjects who were writing more than one word in the blanks of the test, or adding punctuation.

Three words, "injun" (a verb, meaning "to creep up on someone like an Indian"), "nitroglycerine", and "canopies" were defined on the blackboard. In addition, each class was asked if they were familiar with "pheasants" and "partridge." Not one negative reply resulted. During the administration of the test, the investigator was asked for explanations of the following terms: "wander," "undeniable," and "suspenders." All of these requests were from grade seven classes. After completion of the test, only three subjects expressed "having trouble reading the test." This response caused the examiner to question teachers and re-examine school records for evidence of reading difficulties, a search which brought about the decision to eliminate four subjects from the sample population. (See discussion under "The Sample.")

All subjects finished in ample time, one grade seven student taking only 22 minutes. Subjects were encouraged to re-read answers to make certain that they "made sense."

D. SCORING

The literature indicates three ways of scoring the "Cloze" test:

1. acceptance of only the exact word deleted,
2. acceptance of synonyms as well, and
3. acceptance of any word representing the same part of speech as the original.

The criterion adopted for this study is somewhat like number three above. Because the study was seeking evidence that the subjects were capable of using certain structures of language correctly, it was decided that any response or deletion which resulted in reconstruction of the particular grammatical structure incorporated would be judged acceptable, and all others, even if they made perfect sense, would be rejected. This basis of judgment was primarily structural, or at least as structural as is possible when one is investigating language. A problem which might arise in this respect may be illustrated using the acceptable sentence, "John admires honesty," and the ungrammatical sentence, "Honesty admires John." If the first word of each of these sentences were replaced by a blank, one would say that "John" in the first sentence typifies a class of items which would have the features +N, +AN, +M, +HUM Another candidate would be the class +PN, +AN, ±M, +HUM ("He admires honesty", or "She admires honesty.") The second sentence is ungrammatical because "honesty" is inanimate, abstract, and cannot "admire" anyone or anything. Consequently, the criteria of acceptability of answers to the test of this study could not be purely structural; one of the instructions to subjects was that the sentences must "make sense," (i.e., "be grammatical"). Consequently, in the examples above, we cannot say that the insertion of any noun in example one would be acceptable. A requirement of grammaticality, in the sense of being acceptable

as a sentence in English by a native speaker, is assumed to be overriding. Such a criterion is dependent upon a correspondence of feature matrices of items, and inasmuch as such features are semantic, so are the criteria of acceptability of insertions for the instrument of this study.

Whether a scoring criterion of accepting only the exact word deleted would have greatly altered the correlations of this study is an interesting question for replication research. Certainly such a practice would greatly ease the marking burden of such tests.

Student Self-Rating

One small addition to the instruction sheet of the instrument was an item asking students to rate their own ability in English from Poor (value 1) to Very Good (value 5). While not an essential part of the present study, correlation of such self-ratings against syntactic ability scores and writing scores was believed capable of providing some potentially interesting additional findings related to self-concept versus reality.

CHAPTER V

THE WRITING ASSIGNMENTS

In addition to the factors which evaluation creates, the entire writing process for school children is beset by innumerable influences which any examiner finds difficult to control or compensate for. In short, collecting a valid sample of writing from students is a hard task. Further, it is not made any easier if, in order to give extra time for writing, gymnasium periods are taken away. In addition, students are suspicious and hesitant about putting their thoughts on paper for a stranger to read. An investigator can only hope that such factors will average out, and that the relative merits of the samples he collects will remain in their true relationship. There are, however, some factors which an investigator can attempt to control.

A. PHYSICAL SURROUNDINGS

Samples were collected in the morning when it was hoped that students would be well rested and alert. Their regular English or language arts classes were extended somewhat in order to get a sufficient sample from all subjects. The rooms used were their regular classrooms, and normal seating arrangements were followed. The assignments were

written carefully on the chalkboard, and each assignment was carefully explained as well by the investigator. Virtually no distractions or interruptions occurred during the writing periods. When subjects finished their writing, they went on with reading or other work, and did not disturb their neighbors.

B. TIME AND LENGTH OF ASSIGNMENTS

One hour was given for each writing task. Students were asked to aim at 250 words in that time, and were told not to worry about small errors in spelling, mechanics, or penmanship. They were told that recopying was not necessary, so they need not waste time on it.

C. THE SAMPLE SIZE

The original number of students in the experimental group totalled 149. A number did not write one of the writing assignments, or the "cloze" test, and were eliminated. A further elimination of students whose records showed reading difficulty resulted in the final sample having 45 students in grade seven, 49 in grade nine, and 40 in grade eleven, a total N of 134.

A number of research studies have indicated that one writing sample is not sufficient evidence upon which to judge

writing ability.^{1,2} However Diederich,³ in examining the results of University of Chicago students who wrote second and third test essays, noted that one-quarter of the students changed their marks by means of second sample, while less than five percent subsequently changed their marks on the third sample. (While such results would seem to indicate that two samples are adequate, Diederich has apparently since declared two samples to be "totally inadequate.") The investigator has chosen to consider the advantages of a third sample to be questionable, and thus decided to collect two samples of writing from every subject.

D. TOPICS

Choosing a topic on which all students in grades seven, nine, and eleven can write is a problematical task. The investigator was abundantly aware of the enormous effect that topic choice can have on students. However, he was also aware that open-ended topics which would result in completely

¹Gerald L. Kincaid, "Some Factors Affecting Variations in the Quality of Students' Writing." (Unpublished Ed. D. dissertation, [Michigan State College] Michigan State University), 1953.

²C. C. Anderson, "The New S.T.E.P. Essay Test as a Measure of Composition Ability," Educational and Psychological Measurement, XX (Spring, 1960), pp. 95 - 102.

³Paul Diederich, "The Measurement of Skill in Writing," School Review, LIV (December, 1946), pp. 586, 587.

different modes of discourse being taken by some subjects would create great problems for markers. Consequently, it seemed necessary to pick topics which everyone could write on, but which would also control mode of discourse. In order to get some variety of mode, the investigator finally picked one topic designed to require narrative-descriptive writing, and which would permit possible use of dialogue, and a second topic designed to elicit expository-descriptive-argumentive writing. A total of six topics were prepared in each category, and students not in the experimental group were asked to comment on whether they could write freely on any or all of them. Responses were not uniform (as one might expect); but the two topics finally chosen, while not always chosen at the top of each list, received consistent positive response. They are the following:

Topic 1. Narrative:

"It was the worst situation he had ever been in!"

she

they

:

(whatever)

Topic 2. Expository:

"What it means to live in Canada."

It was also felt that topic 1 would permit students to be somewhat imaginative, while topic 2 would get a more pedestrian, consistent reaction. (In point of fact, students

predictably groaned at both.) After a short period of time, the great majority of students seemed to settle down and work quite consistently. The few who "seemed stuck" were given one or two prods, and finally got to work.

Summary

Two writing samples were collected from each of 134 subjects in grades seven, nine, and eleven. Students wrote for one hour on each occasion on an assigned topic. Samples were collected unrevised.

CHAPTER VI

DESIGN OF THE STUDY AND PROCEDURES

A. INTRODUCTION

This introduction outlines the general design of the study. Specific sections to follow discuss the experimental subjects and the sample, the design of the tests, test administration, the statistical analyses and treatment of data, and the significance level.

The initial part of this study was devised to investigate whether a "cloze and delete" test could distinguish between the relative difficulty of certain transformationally-defined structures of English syntax. The test instrument incorporated three occurrences each of thirty-eight such structures. In the "cloze" portion of the test, key words were deleted from the incorporated structures themselves, rather than using the normal "cloze" procedure of deleting every nth word. In the "delete" portion of the test, portions of sentences which included deletable items were underlined, and subjects were asked to cross out or delete as much as they could ("without changing meaning") from the underlined portions. It was reasoned that three occurrences of each of the test structures would provide a variety of contextual clues, thus decreasing the possibility of a lack of response because of unfamiliar

context alone. It was further felt that a greater number of occurrences would make the test too long. Because each structure appeared exactly the same number of times, each subject would have exactly the same number of chances of responding correctly to each item. Consequently, with the final number of subjects being 134, each item was exposed to 402 responses. A straight percentage of correct responses was calculated to measure relative difficulty of structures as indicated by response to this test instrument by the test population.

Because the investigator wished to examine the relationship between subjects' relative syntactic ability, as indicated by the test above, and their writing ability, two writing samples were elicited from each subject. All subjects wrote on similar topics for one hour each. Topics were devised to elicit both narrative and expository writing. It proved impossible to control for length of written themes. All classes wrote for the same length of time and in the same or very similar physical surroundings. Only the investigator was present with the subject classes at the time of writing. The writing samples were thoroughly randomized across grades, numbered, and re-typed exactly as written. The two groups of papers were kept intact as to topic of writing. Each group was marked by three markers who worked independently. Papers were sorted into nine piles, and eventually into five piles of equal number by each marker, who then assigned grades ranging from 1 to 5. Scores for each child were

added to get a composite score, which might range from 6 to 30.

The third part of the study consisted of correlating writing scores and "syntactic ability" scores. This latter score was obtained in the following manner: each transformation was given a "difficulty index score," obtained by subtracting the total percentage of correct responses to it in part one from unity. Thus a scale of weighted scores was obtained for all of the structures of the "cloze and delete" test. Each time a subject replied acceptably to a test item, a number equal to the "difficulty index" for that item was added to his total "syntactic ability score." Finally, "syntactic ability" scores and writing scores for each student were correlated using the Pearson Product Moment approach.

In the final portion of this investigation, the data was examined for evidence of developmental trends in both writing ability and syntactic ability.

Data regarding the subjects' ages, intelligence quotients, sex, and "self rating in English ability" was also collected, and while examination of such data does not constitute a part of the study proper, it will be presented in appendices, and discussed in the final chapter.

B. THE SUBJECTS AND THE SAMPLE

The test population was drawn from pupils in three schools in the City of Edmonton assigned to the investigator by the Edmonton Public School Board. School 1 was a combined Elementary-Junior High School in an area of average to high average socio-economic status. School 2 was a Junior High School in a low to average socio-economic status area. School 3 was a high school of mixed socio-economic makeup. A grade seven and a grade nine class were drawn from each of schools 1 and 2, and two grade eleven classes termed "average" by their teachers were drawn from school 3. Because the principals of school 2 refused to permit any socio-economic questionnaire to be administered, no examination of the effect upon language ability of that particular dimension was possible.

A further objection to the administration of tests of mental ability was raised by the principals of two of the schools on the grounds that standardized tests had been administered recently to the pupils in their schools. Consequently, the investigator was unable to administer such tests. Lorge-Thorndike verbal intelligence quotients were available from the cumulative records of all subjects in the investigation, and the relationship of that factor to the results of the present study is discussed in the final chapter; however, no specific hypotheses are set forth due to a lack of original data. For the sake of comparison,

however, the following data is given:

TABLE 1

Lorge-Thorndike, Form 4, Intelligence Quotient Data-Verbal

| | Range | Mean |
|--------------|--------|--------|
| Grade seven | 84-147 | 114.07 |
| Grade nine | 85-145 | 109.86 |
| Grade eleven | 95-157 | 119.10 |
| Overall mean | | 114.02 |

Ages of subjects ranged from eleven years, seven months to fourteen years, three months in grade seven; from fourteen years, two months to sixteen years, three months in grade nine, and from fifteen years, four months to eighteen years, one month in grade eleven, a total range of six years, six months. The overall average was 177 months, or fourteen years, nine months.

The sample was assigned to the investigator completely by chance; consequently, the final group of subjects consisted of 74 girls and 60 boys. The numbers in the various classes are given in the following table.

TABLE 2

Number of Subjects by Grade and Sex

| | | | Total per Grade |
|--------------|---|----|-----------------|
| Grade Seven | M | 16 | |
| | F | 29 | 45 |
| Grade Nine | M | 26 | |
| | F | 23 | 49 |
| Grade Eleven | M | 23 | |
| | F | 17 | 40 |
| Total N | | | 134 |

Because of the design of this study, it was necessary for every subject to complete all three parts -- the "cloze-delete" test, and two writing assignments -- to be included in the sample. A number of students were thus eliminated.

One further potential problem caused four additional students to be eliminated. Should a subject be a poor reader -- that is, if he should be unable to read the test materials themselves -- his low score would reflect, not poor syntactic ability, but merely his reading weakness. A consultation with classroom teachers and a search of each subject's cumulative record revealed four students with such problems. They were eliminated from the experiment.

Thus the final sample was heterogeneous in nature, and, while slightly above the system average in intelligence,

might be considered average. One slight discrepancy was caused at the grade eleven level, however. Students of lower I.Q. level - from 85 - 95 - seem to have been eliminated, probably being sent to technical or vocational high schools. However, the grade seven and nine samples of the study contained only five pupils in this category, and three of them were in the 91 - 94 range.

C. TEST DESIGN

Four factors influenced the writing of the passages which were to become the vehicle for the "cloze and delete" test:

1. the desire to use a number of modes of discourse, in order to provide variety and interest, as well as to make the use of some of the test structures more natural (it is difficult to use dialogue in exposition).
2. the knowledge that a variety of topics would help to eliminate the possibility of a student's reacting badly toward, or becoming bored with a single topic,
3. the intention to keep vocabulary as simple as possible in order to avoid semantic difficulties, but also,
4. the hope that the passages themselves might be interesting enough that a subject might want to read each passage through, and understand or comprehend it, or at least get something out of it.

A number of passages in various modes were written and submitted to students in the test age range for reaction, the criteria being, "Would you enjoy reading this?" "Does this give you any ideas for writing of your own?" and "Could you react to this in some way?". Narrative selections and descriptive selections that people enjoyed reading, and which gave them ideas for their own writing were judged superior, while expository or argumentative passages that got the strongest reaction, whether positive or negative, were judged most effective.

The rationale for and reaction to each of the passages finally chosen are given below:

Selection 1. Charlie's Chore

Is there a child alive who wouldn't rather get an unpleasant task over with at one swoop? And even if disaster strikes, won't they do exactly the same thing the next time? This little narrative was uniformly well received, and gave a number of students ideas for "human nature" stories of their own.

Selection 2. My Time of Year

Descriptions of places and events seemed to get lukewarm reactions from students who read them, possibly because of a lack of personal significance. This selection seemed to get a better reaction because every student who read it could react to it in some way: we have all experienced the progression of the seasons and we all have our favorite

one for various reasons. For some reason, most students questioned favored summer, although there were several avid skiers as well.

Selection 3. Old Shorty

Whether it was a result of the "back to nature" movement, or whether students just like individualism of any kind, they liked Old Shorty. The model for this character study was a real person, and the investigator liked him too. However, several students thought his philosophy . . . "sort of dumb. . . ."

Selection 4. Peace in Our Time

The reaction to this selection was surprisingly strong. Teenagers seem far more concerned about whether there will be a tomorrow than some of their day-to-day actions might indicate. At one point, a lively discussion on ideologies and why peace would be difficult to achieve grew out of a discussion of this story. Most students rejected the "greater tolerance" solution presented in the passage itself.

After an initial "reaction period," the four passages were chosen, and then re-written somewhat so that all of the test items were incorporated. Reaction to these written passages did not differ from that toward the earlier versions.

The deletions and underlinings were made in the following manner: whenever possible, a function word rather

than a form or content word was chosen for deletion, but the word chosen had to form a necessary part or perform a specific function within the transformation being examined. Several illustrations from "Peace in Our Time" will illustrate, the word in brackets being the word left out.

. . . The (1) (fact) that conflict never seems to cease is undeniable. (2) (What) is the answer? (3) (To) solve this riddle is a [n] (4) (absolute) necessity because of modern technology. Never (5) (before) had man been capable of destroying (6) (himself) and the rest of the world as well . . .

The instructions to students were that they could put only one word, or a contraction such as "can't" or "don't," in a blank. They were to put in the word which they thought "would fit best". Blank number 1 is looking for a factive: no other construction will fit properly. Number 2 demands a WH question, the word "what" being best, but "where" being acceptable. Blank number 3 appears in a "for-to complement" structure: only the word "to" will do the job. (A number of subjects put in the word "we," probably thinking of "That we solve this riddle . . .") Blank 4 demands something modifying "necessity," a pre-nominal adjective of some kind. Any of the following -- and others -- would have been acceptable: great, huge, important, overpowering, real, genuine, -- even "necessary." The criterion of acceptance was whether the word supplied re-created the underlying transformation. From the same article, the following passage illustrates:

The nations of the world find themselves in a strange situation with respect to the promotion of peace. The powerful nations do (7) (not) trust one another, (8) (and) the weaker nations trust nobody. . .

Pilot testing indicated that sufficient context clues existed to make it clear that a negative was wanted in blank number 7. Consequently, if a subject supplied "always," his answer was considered unacceptable. That is, he had not re-created the required negative construction. That his answer was good (but overly naive) English was immaterial.

Delete items were constructed as follows: portions of sentences were underlined, and the subjects were asked to "cross out" as many words as possible without losing any meaning from the passage.

"Show me a city man," said (Shorty) often, "and (9) I'll show you a man who is breathin' poison! / (DELETE) He often (10) said that he would sooner be a hermit (DELETE) than have to live forever in a city.

The acceptable response in delete item number nine would be to strike out "who is," an example of a "WH-be" deletion. Number ten calls for deletion of "that," a complement or "complementizer" deletion. Subjects were told that there were no trick questions. All underlined passages contained at least one word which was to be crossed out.

One worry of the investigator was as to the equivalence of "Cloze" items and "delete" items. A "Cloze" item seems to require more on the part of the subject. He must pull his answer almost from out of the air, guided by context and his knowledge of structure. Consequently, the correct completion of "Cloze" items is very much akin to a

creative endeavour. The "delete" item, however, is very much simpler; there are only so many potential right responses. One can imagine the subject reading the entire sentence, deciding that the word "that" is a candidate for deletion, but also looking at the word "sooner." As he read the entire sentence, however, he would see the necessity for "sooner." He might even recognize that "have to" could have been eliminated if it had been underlined. The process, however, is one of elimination rather than one of creation. Whether these items did prove comparable will be discussed in the chapter dealing with the results of the study.

C. TEST ADMINISTRATION, WRITING SAMPLE
COLLECTION, AND MARKING

The time schedule for administration of the "cloze and delete" test and collection of writing samples was as follows: all subjects wrote on Topic 1, "It was the worst situation () had ever been in," on the morning of the first day they met with the investigator. After one intervening day, they wrote the "cloze and delete" test, and finally, with another day intervening, they wrote on, "What It Means to Live in Canada." This procedure was followed in order to permit a "cooling off" period, particularly for students who might be experiencing some apprehension in what would seem to many of them to be an examination situation. The "before and after" design with the "cloze and delete" test appearing in the middle of the writing assignments was

used so that it could be asserted that whatever they did on the test might safely be attributed at least in part to the same skills evidenced in their writing. Because of the close proximity in time of all three treatments, maturational differences might be safely declared negligible. Any potential learning effects attributable to the test itself would be experienced by all subjects, and would appear in their second set of themes. In effect, the treatment for all subjects would thus remain identical.

The instructions for the "complete and delete test," as it was called, were as follows:

COMPLETE AND DELETE TEST

You are asked to help us in the writing of several stories as follows:

INSTRUCTIONS:

This test involves two procedures:

1. Filling in blanks with SINGLE words so that the sentences fit into the passages or stories and make sense.
2. From the passages or parts of sentences that are underlined, crossing out as many words as you can without changing the sense or meaning of that sentence.

EXAMPLES OF COMPLETION ITEMS:

1. A man shot the dog.

(If you wrote any one word in the blank which described the dog, your answer would be acceptable. e.g. big, black, barking, savage, sick, etc.)

2. I saw the man the moon.

(The obvious word in the above blank would be "in," but other words would also fit: from, with, throwing, etc.)

EXAMPLES OF DELETION (CROSS OUT) ITEMS:

1. John eats grapes because he like to eat grapes. (DELETE)
 (If you crossed out the second "eat grapes," your answer would be acceptable. Cross out unnecessary words in either of two ways, either

...he likes to ~~eat grapes.~~/

or ...he likes to ~~eat grapes.~~/

2. Sam saw an ice cream truck which was in the park./ (DELETE)

(If you crossed out "which was", your answer would be acceptable. Be sure that you cross out everything you want to cross out, and only what you want to cross out.)

NOTE:

*Observe PUNCTUATION, and DO NOT add any of your own.

*Read every sentence completely before filling in blanks. Put in the word that fits BEST and MOST NATURALLY for you.

*Work quickly. Try all test items. You can come back if necessary.

Subjects were reminded about using only one word in a blank, and were reminded to re-read every sentence to see if what they had written made sense.

As outlined earlier, words judged to be at all difficult or unusual in the test were written on the chalkboard, and defined. In addition, students were asked to raise their hands for assistance from the investigator if they should not understand the meaning of any word or expression.

Once the "complete and delete test" was under way, the investigator quickly examined the first few answers of all subjects in order to determine whether they were

following instructions. Once under way, students had few problems, and completed the test easily within the allotted time.

Topics for writing were written on the chalkboard, and read to each class as well. General instructions were as follows:

I would like to see just how well you can write on this topic in the space of an hour. Take a few minutes to plan what you might say on the scratch-paper provided, and then go ahead and write, aiming at 250 words. You need not re-copy your work. If you make mistakes, or want to change a word, cross out the unwanted part neatly, and write in what you want to say above. Do not worry about spelling mistakes. If you must use a dictionary, go ahead and use one. Please, don't consult with your neighbors about approaches to take. Make this your own work, and the best you can do in the time available. Go ahead with your planning now, and try to give us a good sample of your best work.

In addition, a very minor explication of the topics was added.

The length of responses appeared to be highly variable, although most were of adequate length --at least 200 words.

The markers chosen to evaluate the two sets of papers of the present study were well qualified and experienced as indicated by the following table.

TABLE 3

Qualification and Experience of Markers of the Study

| | Degrees | Years of Training | Teaching Experience Schools | University and Other |
|----------|--|-------------------------|-----------------------------------|----------------------------|
| Marker A | B.A. (Hon.), M.Ed., 2nd year Ph.D. | 8 | 6 | 2 |
| Marker B | M.A., B.Ed., 2nd yr. Ph.D. | 8 | 9 | 2 |
| Marker C | B.A. (Hon.), B.Ed. | 5 | 7 | 2 |

All were interested in the task, and particularly interested in how their grades would agree with those of the other two markers.

Instructions to the markers consisted of outlining the general impression marking method, and describing briefly the criteria which the investigator felt would distinguish a better writer from a poorer one. The general impression marking system was described as involving a quick one-to-two minute reading of a paper in order to get a feel of the writer's overall fluency with, and command of, language as a communication medium. Markers were told to disregard spelling errors, and to try to let variations in length of papers be as little influence as possible. It was made clear that quality should be more important than quantity. The

primary criterion established for evaluation was "the writer's ability to use language as an effective means of communicating ideas." The investigator particularly avoided asking the markers to take note of variety of sentence structure.

Working independently, each marker formed nine piles of themes of progressively better quality, and then consolidated these into five piles of near-equal size. Each paper was given a value ranging from 1 to 5, 1 representing "poor," and 5 representing "superior." Each set was marked at one sitting.

D. TREATMENT OF DATA AND STATISTICAL ANALYSIS TO BE APPLIED

1. Complete and Delete Test:

All scoring was done by the examiner using the criteria outlined under Chapter IV, The Test Instrument, and in Section B of this chapter. A tally of acceptable responses, unacceptable responses, and no response was kept according to transformation and subject. Thus individual scores, grade scores, and transformation scores were available. The means and standard deviation of scores for each class and the entire group of subjects were calculated.

The totals of correct responses to each transformation were used to prepare a table and graph of "relative difficulty." Then, transformations were grouped into the following groups: embedding, simple, conjoining, position shift, and

deletion. A table and graph were prepared showing relative difficulty of transformations within groups. Also a composite graph was prepared superimposing the five groups.

A one-way analysis of variance was performed on the average of the various groups of transformations for each class for purposes of examining developmental trends. A table was prepared. Also \bar{Z} scores were calculated for each transformation in order to establish the significance of developmental patterns across grades for each transformation. A table was prepared.

2. Writing Samples:

The means and standard deviations of overall and grade writing scores were calculated. A test of generalizability and reliability based upon parallel samples was performed to determine inter-marker reliability for the markers of the study. Calculations were also performed using the generalizability formula in order to determine the possible effect of more writing samples and/or more markers.

3. Cloze/Writing Correlation:

A Pearson Product-Moment correlation between the "complete and delete" test scores and composite writing scores was performed for the entire sample and for each grade.

4. Other Tests:

A further calculation of the inter-effect of ability on the "complete and delete" and writing ability was performed with the effect of I.Q. factored out.

Inter-correlation of cloze score, writing score, I.Q., age, sex, and student's self-rating in English ability were calculated using the Pearson Product Moment method, for the group as a whole and for grades, and tables prepared.

E. SIGNIFICANCE LEVEL

The significance level adopted in most studies of the present nature for acceptance or rejection of the hypotheses has been .05; however, a figure of .01 was used for the developmental effects of the present study.

Inter-marker reliability in the marking of writing has not achieved very high figures in most studies. Test-retest reliability for an individual marker usually does not rise above 0.70 unless very rigid scales for analytic marking are used. Reliabilities for four markers using pooled scores have achieved figures over .90, but such situations have also judged papers on a pass/fail basis, and consequently spreads of scores have not been large. For the present study, using three markers, two samples of writing, and a parallel-test analysis, an inter-marker reliability of .55 would be accepted as satisfactory.

Correlations between "Cloze" test results and I.Q. scores have ranged between .52 and .60 in the literature for

samples ranging in size from 30 to 104. For the sample size of the present study, a correlation of .60 between "Cloze" test ability and I.Q. would compare favourably with other research, while something higher would be hoped for in correlating writing ability scores and "complete and delete" or "Cloze" test scores. Unless this latter correlation is higher than the correlation between writing scores and I.Q., then the worth of the test of the present study would be doubtful. Of particular interest will be the amount of variance unaccounted for when the effects of I.Q. are factored out, leaving a "complete and delete test" and writing score matrix. An intercorrelation of .40, giving a figure of 16% of the variance to attribute to the interaction of these factors would be most satisfactory.

Note: All statistical analyses were performed with the assistance of the Division of Educational Research Services, University of Alberta, using an IBM 360/67 computer.

CHAPTER VII

FINDINGS OF THE STUDY

A. "COMPLETE AND DELETE" AND TRANSFORMATION DIFFICULTY

Mean Scores on "Complete and Delete" Test

A compilation of acceptable responses was made for each individual's "complete and delete" test, and results were totalled for grades and for the entire sample group. Results for individuals within classes are presented in Appendix A. The means and standard deviations for grades and for the entire group are presented in Table 4.

TABLE 4

Means of Scores on "Complete and Delete" Test

| Grade | N | Mean Score | S. D. |
|--------|-----|------------|-------|
| Seven | 45 | 17.725 | 4.59 |
| Nine | 49 | 18.721 | 4.73 |
| Eleven | 40 | 21.812 | 4.25 |
| Total | 134 | 19.309 | 4.81 |

Because of the size of the N's involved, these differences were significant well beyond the .01 level. The sharp rise in mean scores between grades nine

and eleven, a rise of almost 3.1 per subject, is especially interesting.*

Transformation Difficulty as Indicated from Test Scores

The percentage of correct or acceptable responses for each of the transformations incorporated in the "complete and delete" test was calculated, and appears in Table 5. The table has been rearranged so that the "most difficult" transformation appears first. Column 3 gives percentage of correct response, column 4 gives the resulting "difficulty index" or additive scores credited to a pupil for each acceptable response, and column 5 indicates the group of the transformation, E indicating an embedding transformation, S a simple, C a conjoining, PS a position shift, and D a deletion. The numbering of transformations corresponds to the list given in Chapter III, The Grammar of the Study. Note that seven transformations cover the span of difficulty indexes from .400 to .856, fifteen cover the span from .200 to .400, and sixteen cover the span from .047 to .200. In other words, subjects found relatively few structures to be extremely difficult; the majority of structures seem to be easy or of low-medium difficulty, with a number of embedding operations the most difficult.

*NOTE: The effect of differences in mean I.Q.'s must be considered when examining developmental trends. The 5% drop in I.Q. from grade seven to grade nine suggests that the 6% rise in syntactic ability scores might have been greater with equalized groups. The 9% rise in the mean I.Q. from grade nine to grade eleven would offset some of the 17% increase in syntactic ability mean score. All developmental comparisons must be considered in this light.

TABLE 5

Difficulty of the Transformations of the Complete and Delete Test

| No. | Transformation | Acceptable Res. | Difficulty Index | Type of Transformation |
|-----|--------------------------------------|-----------------|------------------|------------------------|
| 13 | Poss Ing Comp. (Obj.) | .144 | .856 | E |
| 12 | (Sub.) | .229 | .771 | E |
| 8 | It-That Comp. (Sub.) | .483 | .517 | E |
| 36 | Imperative | .512 | .488 | S |
| 24 | Adverbial Expansion of Manner + C | .540 | .460 | E |
| 15 | Factive Insert | .542 | .458 | E |
| 16 | WH Complementizer Insert | .567 | .433 | E |
| 4 | Gerundive | .600 | .400 | E |
| 23 | Adverbial Replacement | .617 | .383 | E |
| 10 | For-to Comp. (Sub.) | .619 | .381 | E |
| 11 | For-to Comp. (Obj.) | .669 | .331 | E |
| 17 | If-Whether Insert | .669 | .331 | E |
| 32 | Adverb. Replacement Shift | .682 | .318 | P.S. |
| 31 | Adverb. Position Shift | .709 | .291 | P.S. |
| 21 | Common Elements Del. | .714 | .286 | D |
| 38 | Reflexive | .714 | .286 | S |
| 37 | Pronominalization | .736 | .264 | S |
| 19 | WH Deletion | .739 | .261 | D |
| 33 | Negative | .749 | .251 | S |
| 6 | Genitive | .751 | .249 | E |
| 1 | Relative Clause (Sub.) | .766 | .234 | E |

TABLE 5 (continued)

| No. | Transformation | Acceptable Res. | Difficulty Index | Type of Transformation |
|-----|------------------------|-----------------|------------------|------------------------|
| 26 | Passive | .784 | .216 | S |
| 20 | WH-be Deletion | .796 | .204 | D |
| 28 | Dative Movement | .838 | .162 | PS |
| 25 | Conjunction | .843 | .157 | C |
| 29 | Particle Movement | .851 | .149 | PS |
| 30 | NP-V Inversion | .853 | .147 | PS |
| 22 | Adverbial Replacement | | | |
| | Del. | | | |
| 34 | Yes/No Question | .856 | .144 | D |
| 14 | Complement Deletion | .858 | .142 | S |
| 2 | Relative Clause (obj.) | .861 | .139 | D |
| 9 | In-that Comp. (Obj.) | .863 | .137 | E |
| 3 | Pre-Nominal Adj. | .868 | .132 | E |
| 7 | Appositive | .896 | .104 | E |
| 18 | "That" + S as object | .896 | .104 | E |
| | Deletion (Quote) | .900 | .100 | D |
| 35 | WH. Question | .905 | .095 | S |
| 27 | Here-There Inversion | .923 | .077 | PS |
| 5 | "with" phrase | .953 | .047 | E |

Relative Difficulty of Embedding Transformations by Grades

| No | Grade Seven | | Grade Nine | | Grade Eleven | |
|--------------|--------------|------|--------------|------|--------------|------|
| | Correct Res. | D.I. | Correct Res. | D.I | Correct Res | D.I. |
| 13 | .096 | .904 | .129 | .871 | .217 | .783 |
| 12 | .237 | .763 | .238 | .762 | .408 | .592 |
| 8 | .415 | .585 | .462 | .538 | .583 | .417 |
| 24 | .430 | .570 | .489 | .511 | .683 | .317 |
| 15 | .474 | .526 | .510 | .490 | .700 | .300 |
| 16 | .474 | .526 | .544 | .456 | .700 | .300 |
| 4 | .548 | .452 | .551 | .449 | .700 | .300 |
| 10 | .548 | .452 | .591 | .409 | .700 | .300 |
| 23 | .556 | .444 | .605 | .395 | .717 | .283 |
| 11 | .615 | .385 | .605 | .395 | .733 | .267 |
| 17 | .622 | .378 | .659 | .341 | .742 | .258 |
| 6 | .682 | .318 | .734 | .266 | .817 | .183 |
| 1 | .696 | .304 | .755 | .245 | .850 | .150 |
| 2 | .807 | .193 | .843 | .157 | .858 | .142 |
| 7 | .852 | .148 | .863 | .137 | .858 | .142 |
| 3 | .896 | .104 | .877 | .123 | .908 | .092 |
| 9 | .904 | .096 | .877 | .123 | .933 | .067 |
| 5 | .926 | .074 | .972 | .028 | .958 | .042 |
| Means | .601 | | .628 | | .967 | .033 |
| Overall Mean | | | .656 | | .741 | |

TABLE 7

Relative Difficulty of Simple Transformation by Grades

| No | Grade Seven | | Grade Nine | | Grade Eleven | |
|--------------|--------------|------|-----------------|------|-----------------|------|
| | Correct Res. | D.I. | No Correct Res. | D.I. | No Correct Res. | D.I. |
| 36 | .489 | .511 | 36 | .489 | 36 | .433 |
| 38 | .637 | .363 | 38 | .687 | 37 | .275 |
| 33 | .667 | .333 | 37 | .727 | 33 | .183 |
| 26 | .733 | .267 | 26 | .741 | 38 | .167 |
| 37 | .756 | .244 | 33 | .768 | 26 | .108 |
| 34 | .793 | .207 | 34 | .829 | 34 | .033 |
| 35 | .867 | .133 | 35 | .870 | 35 | .008 |
| Mean | .706 | | | .730 | | .827 |
| Overall Mean | | | | .755 | | |

Relative Difficulty of Position Shift Transformation by Grade

| No | Grade Seven | | Grade Nine | | Grade Eleven | | | |
|--------------|--------------|------|------------|--------------|--------------|----|--------------|------|
| | Correct Res. | D.I. | No | Correct Res. | D.I. | No | Correct Res. | D.I. |
| 32 | .548 | .452 | 32 | .653 | .347 | 31 | .767 | .233 |
| 31 | .637 | .363 | 31 | .727 | .273 | 32 | .867 | .133 |
| 28 | .719 | .281 | 29 | .843 | .157 | 30 | .917 | .083 |
| 29 | .785 | .215 | 30 | .843 | .157 | 29 | .933 | .067 |
| 30 | .859 | .141 | 28 | .857 | .143 | 28 | .950 | .050 |
| 27 | .896 | .104 | 27 | .918 | .082 | 27 | .958 | .042 |
| Mean | .741 | | | .807 | | | .899 | |
| Overall Mean | | | | .815 | | | | |

TABLE 9

Relative Difficulty of Deletion Transformation by Grade

| No | Grade Seven | | Grade Nine | | Grade Eleven | | | |
|--------------|--------------|------|------------|--------------|--------------|----|--------------|------|
| | Correct Res. | D.I. | No | Correct Res. | D.I. | No | Correct Res. | D.I. |
| 21 | .637 | .363 | 19 | .734 | .266 | 19 | .667 | .333 |
| 20 | .778 | .222 | 20 | .748 | .252 | 21 | .733 | .267 |
| 19 | .807 | .193 | 21 | .768 | .232 | 14 | .817 | .183 |
| 22 | .844 | .156 | 22 | .836 | .164 | 20 | .875 | .125 |
| 18 | .882 | .118 | 14 | .857 | .143 | 22 | .892 | .108 |
| 14 | .904 | .096 | 18 | .897 | .103 | 18 | .925 | .075 |
| Mean | .809 | | | .807 | | | .818 | |
| Overall Mean | | | | .811 | | | | |

Transformation Difficulty Within Groups

The preceding tables represent the relative difficulty of various transformations within groups. No table is presented for the conjunction transformation, the difficulty of which for the various classes was as follows: grade seven, acceptable response, .770 (difficulty index .230); grade nine, acceptable response, .836 (index .164); grade eleven, acceptable response .933 (index - .067); overall mean of acceptable response, .846. Note that a comparison of the figure for conjoining with means for the other types of transformations indicates that conjoining is the easiest operation of all, followed by deletion and position shift operation--almost tied and still relatively easy. Then comes the group of simple transformations, and finally, embedding transformation the most difficult type by a fairly wide margin. The data for grades seven, nine, and eleven for the other classes of transformation appears in Tables 6, 7, 8, and 9. Note: graphs of transformation difficulty indices overall and by grades appear in Appendix F.

An analysis of variance of the class means for the four groups having more than one member (i.e.,--the groups other than conjunctions) yielded a significance of difference between groups and of the interaction of groups and grades

well beyond the 0.01 level of significance. Data for the analysis of variance appears in Appendix B.

A Spearman rank order correlation (See Appendix C) was performed on each table above between grades seven and nine, seven and eleven, and nine and eleven in order to determine whether trends in difficulty between grades were consistent for particular transformations: all such correlations for Table 6, embedding transformations, were significant at beyond the .01 level. For simple transformations (Table 7), correlations were significant at the .05 level, but did not reach .01. For position shift (Table 8), only the correlation between grades nine and eleven reached the .05 level. Correlations performed on the data of Table 9 (Deletions) were all not significant.

Consequently, it appears not only that embedding transformations are more difficult as a group than simple transformations, which in turn are more difficult than the other types, but also within groups, both simple and embedding structures (but particularly the latter) retain to a considerable degree their difficulty relative to one another, at least across the school grades represented in this study.

B. THE WRITING ASSIGNMENT AND INTER MARKER RELIABILITY

Mean Scores on Writing Tests

Writing scores for each individual were obtained by adding together six scores: two from each of three

markers for each subject of the study. Raw data appear in Appendix A. As the review of the literature has revealed, a single test of writing does not yield reliable scores; consequently, the following Table does not give means and standard deviations for each writing assignment. The data represents the pooled scores for each subject.

TABLE 10
Means of Scores on Writing Assignments

| Grade | N | Mean Score | S. D. |
|--------|-----|------------|-------|
| Seven | 45 | 16.07 | 5.09 |
| Nine | 49 | 17.04 | 6.28 |
| Eleven | 40 | 21.75 | 5.06 |
| Total | 134 | 18.12 | 6.01 |

Differences in mean scores are all significant beyond the .01 level. As with data on "complete and delete" test scores the wide spread between grades nine and eleven in mean scores, 4.7 points in this case, is especially noticeable. It is possible that the higher than average I.Q. in grade eleven and the slightly lower than average I.Q. in grade nine might account for some of this difference, but not as much difference as seems to exist.*

*See note on page 208.

Inter-Marker Reliability

Rather than using pooled scores to determine inter-marker reliability, the investigator chose to treat the two writing assignments as parallel tests, and to apply an approach to reliability through a generalizability formula suggested by Cronbach, Rajaratman, and Gleser,¹ and further developed for a number of test situations in Gleser, Cronbach, and Rajaratman.² The particular approach chosen was believed suitable for the test situation: the two writing samples could be considered "parallel" tests, and the individual criteria for marking used by the three evaluators would represent the "strata" of treatment. The actual calculations appear in Appendix D; however, the final formula for the particular present study's design is as follows:

$$\text{Reliability} = \frac{\sigma_e^2}{\sigma_e^2 + \frac{1}{T} \sigma_T^2 + \frac{1}{M} \sigma_m^2 + \frac{1}{TxM} \sigma^2}$$

In this formula, T is the number of tests, M the number of markers. Using the figures calculated in Appendix D, the

¹Lee J. Cronbach, Nageswari Rajaratnam, and Goldine C. Gleser, "Theory of Generalizability: A Liberalization of Reliability Theory," The British Journal of Statistical Psychology, XVI (21), (November, 1963), pp. 137-163.

²G. C. Gleser, L. J. Cronbach, and N. Rajaratman, "Generalizability of Stratified Parallel Tests," Psychometrika, XXX, No. 1 (March, 1963), pp. 39-56.

inter-marker reliability of the present study is

$$\frac{.97}{.97 + \frac{.19}{2} + \frac{.29}{3} + \frac{1}{2 \times 3} .97} = 0.73$$

This formula permits a substitution for values of T and M in order to generalize. That is, it is possible to substitute different values directly into the above formula to determine what the effect of 6 markers or only one test would be. For example, six markers would raise inter-marker reliability to 0.82; three markers but only one test would lower the figure to 0.61; three markers and three tests or four markers and two tests (a common arrangement in the literature) would raise the inter-marker reliability to 0.78. For the purposes of the present study, however, the inter-marker reliability achieved (0.73) is considered satisfactory, particularly in the light of the very general criteria for marking given to the evaluators and the lack of a "training program," as is commonly used with the analytical method of marking.

C. PRINCIPAL CORRELATION FOR THE GROUP AS A WHOLE

Pearson product-moment correlation methods were applied to data for the entire population. Table 11 gives a complete correlation matrix for the entire sample of 134 subjects using six variables:

1. Intelligence quotients,
2. Cloze test scores,

3. Sex (1 = male, 2 = female,)
4. Students' self-rating in English, from
1 = poor to 5 = superior,
5. Age in months
6. Total writing score.

While several of these correlations are of interest, the most important one for the principal hypothesis of this study is that between factors 2 and 6 - "Cloze" or "complete and delete" test score, and total writing ability score. Also of interest, although not involved in the hypotheses of the present study, are the correlations between I.Q. and "complete and delete" test scores, and between I.Q. and writing scores. For the method of the present study, the preparation of a structurally designed "complete and delete" test, to represent a potentially new way of predicting writing ability, the correlation between it and writing score must be higher than the correlation between I.Q. and writing score.

As Table 11 shows, this is indeed the case, the correlation between "complete and delete" test scores and composite writing scores being 0.73, while the correlation between I.Q. and writing scores is 0.57.* The correlation between I.Q. and "complete and delete" test scores is 0.64, slightly higher than the average correlation between "Cloze" test scores and I.Q. in the literature.

Note also the relatively lower correlation between

*Note: This is also the case with each grade level. See Section E.

age and "complete and delete" test score (0.19; significant beyond the 0.05 level). Although the difference is not great, it appears as though the mere process of "getting older" has some effect upon a person's writing which is not accounted for by increased syntactic ability. The answer might possibly lie in experience, providing better "content" for older writers.

D. INTERRELATIONSHIP OF CLOZE AND WRITING SCORE WITH
I.Q. FACTORED OUT

Using a method outlined in McNemar,³ the investigator calculated the correlation of "Cloze" (or "complete and delete") test scores interacting with writing ability scores with the effect of I.Q. factored out. The following formula from p. 166 of McNemar states this relationship:

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2} \sqrt{1-r_{23}^2}}$$

Letting subscript 1 be "complete and delete" test scores, 2 be writing scores, and 3 be I.Q.'s, and using the figures from correlationship Table 11, the correlationship becomes:

³Quinn McNemar, Psychological Statistics, 3rd. ed. New York: John R. Wiley and Sons., 1962, p. 166.

TABLE 11

Pearson Product-Moment Correlation Matrix for Six Factors of Test Population

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------------|---|------|------|------|-------|------|
| 1. I. Q. | | 0.64 | 0.21 | 0.42 | -0.04 | 0.57 |
| 2. "Complete" Score | | | 0.19 | 0.47 | 0.19 | 0.73 |
| 3. Sex | | | | 0.05 | -0.20 | 0.30 |
| 4. Self-Rating | | | | | 0.05 | 0.26 |
| 5. Age (months) | | | | | | 0.25 |
| 6. Writing Score | | | | | | |

TABLE 12

T-Values and Probabilities Associated With Them for Above Correlation Matrix

| | Probabilities | | | | | |
|---------------------|---------------|-------|-------|-------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1. I. Q. | | 0.00 | 0.02 | 0.00 | 0.61 | 0.00 |
| 2. "Complete" Score | 9.66 | | 0.03 | 0.00 | 0.03 | 0.00 |
| 3. Sex | 2.43 | 2.26 | | 0.58 | 0.02 | 0.00 |
| 4. Self-Rating | 5.38 | 6.12 | 0.55 | | 0.59 | 0.00 |
| 5. Age (months) | -0.52 | 2.20 | -2.32 | -0.54 | | 0.00 |
| 6. Writing Score | 8.05 | 12.17 | 3.60 | 3.07 | 3.03 | |
| | T-Values | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |

$$\begin{aligned}
 r_{12.3} &= \frac{.73 - (.64)(.57)}{\sqrt{1 - (.64)^2} \sqrt{1 - (.57)^2}} \\
 &= \frac{.3652}{(.7684)(.6314)} \\
 &= 0.58
 \end{aligned}$$

This means that $.58^2$, or 33.6 percent of the variance may be attributed to the interaction of factors involved in both the "complete and delete" test and writing ability.

E. CORRELATION MATRICES FOR ALL FACTORS BY GRADES

Having applied the Pearson product-movement correlation method to six data factors for the entire sample population section C of this chapter, the investigator felt that a similar process might be of interest if applied to individual grade populations. Consequently, matrices were prepared and correlations performed for each grade separately using the same factors. Correlations and significance data are presented for each grade in Tables 13-18. Because of the lower population totals for individual grades, it would not be expected that the correlations for grades would be as large as those for the entire population, and such appears to be the case.

TABLE 15

Pearson Product-Moment Correlation Matrix for Six Factors For Grade Nine

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------------|-----|---|-----|------|------|------|
| 1. I. Q. | | | | | | |
| 2. "Complete" score | .74 | | .38 | .49 | -.51 | .58 |
| 3. Sex | | | .35 | .46 | -.46 | .72 |
| 4. Self-Rating | | | | -.02 | -.07 | .58 |
| 5. Age (months) | | | | | -.43 | .15 |
| 6. Writing Score | | | | | | -.30 |

TABLE 16

T-Values and Probabilities Associated With Them for Above Correlation Matrix

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------------|-------|-------|------|-------|-------|-----|
| 1. I. Q. | | | | | | |
| 2. "Complete" score | 7.45 | .00 | .01 | .00 | .00 | .00 |
| 3. Sex | 2.85 | 2.57 | .01 | .00 | .00 | .00 |
| 4. Self-Rating | 3.87 | 3.57 | -.12 | .90 | .64 | .00 |
| 5. Age (months) | -4.08 | -3.51 | -.47 | -3.24 | .00 | .29 |
| 6. Writing Score | 4.85 | 7.06 | 4.89 | 1.06 | -2.18 | .03 |

TABLE 17
Pearson Product-Moment Correlation Matrix for Six Factors for Grade Eleven

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------------|---|-----|-----|------|------|------|
| 1. I. Q. | | .41 | .01 | .21 | -.52 | .52 |
| 2. "Complete" Score | | | .15 | .37 | -.36 | .60 |
| 3. Sex | | | | -.07 | -.08 | .16 |
| 4. Self-Rating | | | | | -.07 | .34 |
| 5. Age (months) | | | | | | -.27 |
| 6. Writing Score | | | | | | |

TABLE 18
T-Values and Probabilities Associated With Them for Above Correlation Matrix

| | Probabilities | | | | | |
|---------------------|---------------|-------|-------|-------|-------|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1. I. Q. | | <.01 | .94 | .18 | .00 | .00 |
| 2. "Complete" Score | 2.73 | | .35 | .02 | .02 | .00 |
| 3. Sex | 0.08 | 0.94 | | .66 | .63 | .32 |
| 4. Self-Rating | 1.35 | 2.42 | -0.44 | | .65 | .03 |
| 5. Age (months) | -3.75 | -2.36 | -0.48 | -0.46 | | .09 |
| 6. Writing Score | 3.79 | 4.63 | 1.00 | 2.23 | -1.76 | |
| | T-Values | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |

Additional Findings for Grade Seven

The most striking figures in the Grade Seven correlation matrix are the correlations between both "complete and delete" test score and writing score and the age factor, $-.49$ and $-.42$ respectively. These figures would seem to indicate that younger, brighter students who may have been subject to early promotion are also the best writers and best users of language. This thesis is supported by the even larger negative correlation between I.Q. and age, $-.52$. A cause for this phenomenon might be found in the programs of the elementary schools at present, which provide 5 year, 6 year and 7 year programs, resulting in a much larger proportion of brighter young children moving into the junior high schools.

Girls are better than boys in both syntactic ability and writing ability, the respective correlations being $.26$ and $.33$, the latter being significant at the $.02$ level. These are the results of a bipart coding system being used, girls being coded 2 and boys 1. Thus any positive correlation favours girls, while a negative one would favour boys.

Letting these students rate themselves in ability might not be the worst idea in the world. Apparently they have an excellent notion of just how good they are, achieving a correlation between self-rating (on a 5-point scale) and writing ability of $.40$ (significant at the $.01$ level), and a correlation between self-rating and syntactic ability as indicated by the "complete and delete" test of $.63$, significant beyond $.01$.

Additional Findings for Grade Nine

The picture of younger students being the better writers and greater masters of the structure of language is repeated very significantly at the grade nine level. The correlation between age and I.Q. is almost identical, $-.51$. The correlations between age and syntactic ability, and age and writing ability are $-.46$, and $-.30$ respectively, the former significant beyond $.01$, the latter significant at the $.03$ level.

Once again, girls outstrip the boys in both syntactic ability and writing ability, but particularly in writing ability.

Correlation between self-rating and syntactic ability or "complete and delete" test score remains high and significant beyond the $.01$ level ($.46$), but the relationship between writing scores and the students' self-ratings is low ($.15$).

The exceptionally high correlation between I.Q. and "cloze" test score should be noted.

Additional Findings for Grade Eleven

This grade produced the lowest correlations of main effects. While the three correlations involving I.Q., "complete" test score, and writing ability score are all significant at beyond the $.01$ level, the size of the correlations is markedly lower. The correlation between the "complete" test and I.Q. is $.41$, still within the range of

such correlations in the literature.

While the girls seem to remain somewhat superior to the boys in both writing ability and syntactic ability, the gap appears to be closing, with neither correlation figure nearing significance.

These older students seem to be somewhat more realistic than the grade nines about their own ability. Their self rating correlates with writing ability and "complete" test scores respectively at levels of .34 (sig. at .03) and .37 (sig. at .02).

Additional Findings of Trends Across Grades

The relationship of major factors remains fairly consistent across all grade levels examined, although the correlation between I.Q. and "complete" test score in grade eleven is somewhat low. In every case, the "complete" test score correlates with writing ability more highly than I.Q. does, although both sets of correlations are consistently significant at beyond the .01 level.

Age correlates consistently negatively with both writing ability and syntactic ability; clearly, the grade system and promotion system in school guarantees (within the sample population of this study, at least) that the younger a student is within his class, the better he or she will be as a student of language arts. In fact, it is very probably that same ability that got the student where he or she is.

Girls appear to be ahead of boys in both writing ability and syntactic ability throughout the grades studied; however, it appears as though this gap is closing by grade eleven.

While students appear to have reasonably good ability to judge their own inherent syntactic ability, their knowledge of the quality of their writing is far more inconsistent.

F. SUMMARY OF FINDINGS

The preparation of a structurally-constructed "complete and delete" test based upon "Cloze" procedure appears to be completely feasible. The use of a scoring system based upon criteria of whether or not subjects reconstruct transformational configurations incorporated into written passage also seems workable.

Such a test appears to be capable of distinguishing to some extent between the relative difficulty, not only of different particular transformations, but also of different classes of transformations. Indications are that so-called "embedding" transformations are clearly more difficult than other types. "Simple" transformations rank second in difficulty, while "position shift" and "conjoining" transformations appear to be relatively easy, with the latter being easiest. The difficulty of "deletion" transformations, while appearing similar to that of "position shift" operations, is hard to establish because of inconsistent response, and possible design weakness.

The weighting system devised for this study for providing "difficulty indexes" appears not only workable, but highly defensible. If 90 percent of students can reply acceptably to a "Cloze" type item, that item cannot distinguish very clearly between subjects; an item receiving 10 percent accurate or acceptable response is clearly harder, discriminates well, and should receive a higher weighting.

The use of this weighting system in assigning additive scores to students, indicating a factor which we might call "syntactic ability" or "maturity," has produced scores for individuals which seem to indicate a developmental trend in such syntactic ability across grades seven, nine, and eleven, with a marked increase in such ability between grades nine and eleven.

A rank order correlation method indicates that the relative difficulty of various transformations within groups, particularly true of embedding transformations, but also to a lesser extent of simple transformation, remains remarkably consistent across the grade levels of the study.

It is possible to achieve acceptably high levels of inter-marker reliability using three well qualified markers, minimal instructions to such markers, and two writing samples on single topics.

The use of the above design produced a set of composite writing ability scores, the means of which for each grade were significantly different for the grade levels of the study. A developmental pattern emerged, with the grade

eleven mean score again notably higher than the grade nine score.

Correlations of individual writing scores and "syntactic ability" scores produced a correlation ratio of 0.73, significant well beyond the .01 level. Correlations between "syntactic ability" scores and I.Q. figures resulted in an r of .64, slightly higher than similar correlation between "Cloze" procedure scores and I.Q. scores recorded in the literature. A correlation of I.Q. and writing ability scores achieved a value of r of .57.

Factoring out the influence of variations in I.Q. in order to determine the interaction of writing ability scores and "syntactic ability" scores resulted in an r of 0.58, meaning that 33.6 percent of the observed variance is accounted for by such interaction.

Comparison of correlation matrices for six factors for each grade indicated a number of other interesting findings. The greater ability of girls in both writing and syntactic knowledge is more marked in grades seven and nine, but becomes less in grade eleven. The younger students at all grade levels were superior in both writing ability and syntactic ability, a probable result of school promotional practices. Students also appear to have some ability to judge their own syntactic knowledge, but less ability to judge their own writing ability.

CHAPTER VIII

RE-EXAMINATION OF HYPOTHESES

The null hypotheses of Chapter I are restated below, and the evidence for acceptance or rejection is given below.

Null Hypothesis A

Individual student's additive syntactic ability scores will not correlate significantly with their individual total writing scores.

Rejected. Not only is the correlation of these two factors considerable ($r = 0.73$) and highly significant (well beyond 0.01 level), but a similar correlation exists within all grades as well. Had the correlation been limited to embedding and simple transformation scores against writing scores, thus eliminating the inconsistencies existing in calculations for other types of transformations, it is possible that this correlation might have been much higher.

The possibility exists that the "complete and delete" test of this study is nothing but another form of intelligence test, and that the subject's ability to respond is merely a reflection of overall mental ability. However, factoring out the effect of intelligence still left a considerable portion of the observed variance (33.6 percent) which could be accounted for only by the interaction of writing ability and syntactic ability.

Null Hypothesis B

No significant patterns of development in syntactic ability as measured by the instrument of this study will appear across grades seven, nine, and eleven.

Rejected. For sixteen out of eighteen embedding transformations of the study, five out of seven simple transformations, the conjoining transformation, and five out of six position shift transformations, developmental trends were present and significant beyond the 0.01 level. In several cases where reverse developmental patterns seemed to exist, a careful scrutiny of original data indicated that subjects were using patterns of transformation not included in this study, and were, as Hunt¹ suggests, actually exhibiting greater syntactic maturity than the test instrument and techniques were designed to diagnose.

The situation with deletion transformations is peculiar, and is probably a fault of test design. The older students got, the more inclined they were to try to do "something extra" in the way of deletions. Instead of performing what seemed almost too obvious a task (as did most of the grade seven subjects), the grade nines and elevens deleted too much, destroying structures and losing meaning-- as well as credit for syntactic maturity. Consequently, the investigator believes that ignoring the deletion transformations in rejecting the above null hypothesis is defensible.

¹Hunt, op. cit., 1970, pp. 53-54.

Null Hypothesis C

No significant patterns of development in writing ability as determined by the methods of this study will appear across grades seven, nine, and eleven.

Rejected. Analysis of variance of mean writing scores for classes clearly shows developmental trends significant beyond the 0.01 level. One problem with such a conclusion, however, is the assumption of similarity of content within the writing samples across grades. While it is true that all subjects wrote on the same two topics, it is undeniably true that every paper would have differed from every other paper in what was done within the bounds set by the topics. In this connection; Hunt² found superior writers managing to alter content even when content was rigidly controlled.

Null Hypothesis D

No significant pattern of development in ability with the individual transformations of this study (embedding, simple, position shift, and deletion) will appear across grades seven, nine, and eleven, and between transformations in groups.

Rejected. When transformations were ranked within groups and grades, a Spearman rank order correlation indicated a consistency in ranking of transformations across grades which reached a level of significance beyond 0.01 for embedding transformations, and beyond 0.05 for simple transformations. Only the relationship of ranking of transformations between grades nine and eleven reached the 0.05 level of significance for position shift transformations, where the low number of

² Ibid., p. 54.

transformations (6) and the overall similarity in difficulty made any significant finding unlikely. The relationship of deletion transformations, for the reasons stated under Null Hypothesis B, was not considered in rejecting this hypothesis.

CHAPTER IX

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

A. SUMMARY OF THE STUDY

This study began with three purposes in mind:

1. To investigate whether it would be possible in some quantifiable way to determine the relative difficulty of some of the transformationally-defined structures of English,
2. To determine whether students of different ages would demonstrate developmental trends in their ability to use the transformational structures investigated above, and
3. To ascertain whether there would be any correlation between the above students' syntactic ability and their ability in written discourse.

A survey of the literature revealed that approaches to identifying what made language "difficult" were mostly unsatisfactory. Work in readability tended to emphasize the mere counting of elements or words, and vocabulary difficulty. One promising approach which seemed to take context and major structures of language into consideration was the "Cloze" procedure, whereby subjects were asked to fill in words deleted from passages. "Cloze" technique was thus adopted for the instrument of this study.

Traditional and structural grammars did not provide the unifying theory and nomenclature desired in the design of this study; however generative-transformational grammar did provide exactly such a framework. Of particular interest were the "classes of transformations" which such theory provided.

Psycholinguistics suggested that there appeared to be some reality to the beliefs that people segment language, that such segmentation aids comprehension, that people process language on several "levels", and that people "processed" language in some fashion corresponding to transformational processes. Further, research indicated, with some clarity, that certain structures of language were of varying degrees of difficulty, although whether that difficulty was determined by transformational history or simply the number of elementary operations involved in their production was not made clear.

Other investigators (and particularly educators), in trying to identify what made certain kinds of writing "more mature" or in some sense "better" than other writing, developed a number of indexes: sentence length, ratio of subordinate to main clauses, length and type of clauses, and finally, length of "communication unit" or T-unit (meaning "terminable unit"). Further investigation into the cause of the greater length of "more mature" T-units showed that processes of sentence combining (much like transformational "embedding" of one sentence into another)

accounted for greater T-unit length, and the r i c h e r highly descriptive quality of better writing.

Several attempts at teaching transformational theory and sentence-combining and lengthening techniques have shown promising results.

In order to fulfill the intentions set out above, the investigator prepared a structurally constructed "complete and delete" test incorporating each of 38 transformational structures three times within four passages of various modes. As in "Cloze" procedure, words were deleted from these transformational structures, but in such a way that, should a subject be able to put an acceptable word in the blank, it would indicate a familiarity with the particular structure incorporated within that context. In order to examine whether students could perform various transformational deletions, a number of sentence parts were underlined, the subjects being asked to delete or cross out as many words as they could without changing meaning. Responses to both types of items were judged acceptable if whatever the subject wrote or deleted corresponded to the particular transformational structure incorporated.

Subjects also wrote two short themes each, the themes to be evaluated holistically by three experienced markers, and placed into five categories of quality. The themes were on two set topics. The student's "writing score" was to consist of the combined marks from all three markers on the two themes.

The sample population consisted of students from three schools in the City of Edmonton: 45 grade seven students, 49 grade nine students, and 40 grade elevens. The total test population was 134.

Total frequency of acceptable responses on the "complete and delete" test was used to determine the relative difficulty of the transformational structures incorporated in the test. The percentage of acceptable responses was subtracted from unity to yield a "difficulty index" for each transformation, and then each subject's "syntactic ability" was determined by crediting him with the appropriate "difficulty index" figure for each acceptable response on his "complete and delete" test.

A correlation of each student's syntactic ability score and his total writing score was performed using the Pearson product-moment method. The effect of I.Q. was also factored out.

Transformations were grouped into classes of embedding, simple (or single-based), conjoining, position shift, and deletion, and classes were compared for significant differences in difficulty using a one-way analysis of variance. Also, the consistency of difficulty ranking of transformations within groups and across grades was computed using a Spearman rank order correlation approach.

Developmental trends across grades for individual transformations were calculated using a method based upon a difference between uncorrelated proportions to give a

standard score.

Finally, correlation matrixes for factors of syntactic ability, writing ability, I.Q., age, sex, and self-rating of ability in English were prepared for each grade.

B. FINDINGS AND CONCLUSIONS

The Difficulty of Transformational Structures

Clearly, the results of this study indicate a great variation in the frequency with which subjects responded accurately or acceptably to the items incorporated in the "complete and delete" test. Subjects were not confused by the task they were attempting, nor did any of them simply give up. Of 15,276 potential responses, only 401 blanks were left, a figure of only 2.7 percent for "failures to respond." Even the slowest workers tried every item.

One important question concerns the underlying assumption of the design of the test: does an accurate response mean that a subject has mastered or recognizes the underlying transformational structures? Conversely, does a failure to respond acceptably clearly indicate a lack of mastery of a particular structure? In marking the individual "complete and delete" tests, the investigator was struck by a particular pattern of response: when a subject failed to recognize a particular structure once, he would quite often miss it again, or even all three times. Further, this was not the case for difficult transformations alone. A number of times, a subject missing a simple position shift

or deletion item (the accurate response for both of which was over 81 percent) would fail to get the next two similar items as well.

At one point the investigator considered an "all or nothing" marking approach, which had two versions: first, demanding accurate response to all three items in order to give credit for a particular transformation; or second, giving credit for a transformational structure if even one acceptable response was received, this latter method based upon the idea, "If he can do it once, he can do it again." Both of these methods were rejected on the grounds that there are degrees of mastery of everything. A person who can complete a structure in one context but not in two more obviously has not mastered that structure as well as someone who can get two out of three, and so on. Consequently, each blank was considered separately in giving credit to individuals for syntactic maturity, and in tallying total acceptable responses in order to determine the scale of difficulty of transformations.

In answer to the question of whether responses to the test items clearly indicate syntactic ability in subjects, the investigator would be inclined to give a tentative "yes." The only certain way of answering this question would be to examine a large corpus of writing from the subjects of the test in order to determine whether they do, in fact, use such structures in their written discourse. And even then, there is no sure way of knowing if a particular structure which did

not appear in 10,000 running words of writing by a subject might not have appeared in the next sentence. However, such a study should be attempted.

An examination of the entire array of transformational structures incorporated in the "complete and delete" test arranged in order of difficulty reveals several which seem to be extremely difficult indeed: number 13, the "possive-ing" complement structure in the object position, and the same structure in the subject position (number 12). Instead of writing, "Autumn's painting of the hillside was beautiful," subjects would fill in, "Autumn painting of the hillside . . ." Another example would be, "There is no doubt that the farmer's hardest toil of the year is represented by his autumn harvesting." Subjects would ignore the parallelism with the earlier possive-ing structure and fill in " . . . is represented by late autumn harvesting." or some such. Item number 13 was responded to correctly by only 9.6 percent of the grade sevens, 12.9 percent of grade nines, and 21.7 percent of grade elevens. One is tempted to speculate that this particular structure is disappearing from common English usage. However, it is also interesting to note that most of the acceptable responses to these items were submitted by subjects whose overall response was exceptionally good. The exceptionally high weighting which these items received (particularly item number 13) was somewhat unfortunate, and might have distorted overall scores somewhat. Item number 8 is also a complement structure, and an embedding. Its

position as third on the list of difficulty is not surprising. A sentence of the form, "That Joe is an idiot is obvious," is completely foreign to students, who would say, "It is obvious that Joe is an idiot." This last structure is number 9, the same it-that complement, but in the object position. Its difficulty index in the subject position is .517; in the object position, it is .139, placing item number 9 seventh from the last in the list of difficulty. Obviously, where items appear in sentences has an important effect upon their difficulty, one factor which the present study has probably not taken sufficiently into account.

Of the first twelve most difficult transformational structures, eleven fall into the "embedding" category. The lone intruder, ranking fourth, is the imperative, a simple transformational structure. About the only logical way to prepare a "structurally constructed" cloze item for the imperative structure is to delete the verb. The following is an example from the test: "He always said, 'Imagine you're always workin' for yourself, and then don't disappoint the boss!'" Even with the parallel imperative in the second clause, a great number of subjects could not get this item. Perhaps such items lack some contextual clues. On the surface, at least, the imperative does not seem to be that difficult.

The investigator was astonished at the number of mistakes made with the simple process of conjunction, particularly at the grade seven level. By grade eleven, only

6.7 percent of the responses for the conjunction structure were unacceptable; but at the grade seven level, 23 percent of responses were wrong. With 16.4 percent of grade nines' responses being unacceptable, a clear developmental trend and a most unexpected one, appears. Most investigators seem to think that the process of conjunction is mastered early. This study indicates that, while young children may use a lot of conjunctions, they don't seem to use them accurately until about grade eleven.

A number of embedding structures appear among the eight easiest transformations. Relative clauses (number 2) and "it-that" complement structures (number 9) are easy in the objective position. Pre-nominal adjectives and appositives (numbers 3 and 7 respectively) are even easier, the latter somewhat surprisingly, as several studies have indicated appositives to be difficult. One answer to this situation might be that while the use of appositives may be somewhat uncommon among young writers, it is not in itself a difficult structure, and is easily recognized because of the abundance of contextual clues which surround it. The easiest transformation of all, the "with phrase," (number 5), had an incorrect response rate ranging between 7.4 and 2.8 percent. The investigator expected to get "A bag having a wet bottom . . . ", but everyone seemed satisfied with "with."

The overall picture which emerges with respect to the relative difficulty of the various structures incorporated in the test instrument of this study is very much like that

reported in studies by Loban,¹ Hunt,² O'Donnell,³ and Fagan.⁴ The "easiest" structures seem to be those in common conversational usage, while the most difficult ones are typical of written language. What would be the reaction of a friend if you turned to him at a football game and said, "That the quarterback is a stumble-bum is obvious"? The reaction you would receive is even more obvious. People simply don't talk in that fashion; the influence of patterns of spoken language upon written language is overwhelmingly greater than the converse, particularly -- as the present study and others indicate -- with younger children -- or adolescents.

The present study has examined the reactions of a very limited group of subjects to an extremely limited corpus of language. The technique, however, has forced these subjects to use, or manipulate, or in some fashion come to grips with certain structures of language. The entire process has dealt with written language, however, and any speculations with respect to the influence of the experiential saturation of oral or spoken language which humans undergo daily cannot be empirically based. Several rather simple logical inferences can be drawn though, which seem to bear upon the relationship between the "difficulty" of some structures of language and the efforts of educators

¹Loban, op. cit.

²Hunt, op. cit.

³O'Donnell, op. cit.

⁴Fagan, op. cit.

to teach youngsters to "write better."

First, it is clear that what most critics identify as "skilled", or "mature" writing has somewhat less relationship to the functional, direct, and simple patterns of speech than might at first be expected. Good writing is not good speech written down. The content may be the same, and the logical development of the content may be similar; but the patterns of language used will show striking differences.

We do not write in the same fashion-- using the same syntactic structures-- as we talk, although adopting the simple patterns of speech into writing is accepted, but often branded "immature." Using the patterns of written language in speech is a practice limited to badly written speeches, plays, movies, and television shows. The surest indicator of a bad dramatic performance (which supposedly seeks realism) is a negative answer to the question, "Do people really talk like that?" What would be the reaction of most people if, in reply to the District Attorney's question, "Miss Jones, where were you on the night of July tenth at 10 P.M?," she replied, "Tired, hungry and full of despair, I went home, eager for a tub of hot water, soothing, cleansing, bubbly, its warmth easing every muscle, its"

Surely even the most avid supporter of the "cumulative sentence" would recognize its ludicrousness in oral language. Such structures have their place. They can perform masterfully in descriptive phrase, or in sensitive narration.

But there is a great difference between oral and written patterns of language.

Young children do not know the world of written language. Theirs is a world of "oracy," a world of sounds and voices, a world which speaks to them directly, and makes every effort to understand their most inadequate efforts at communication. As the child grows older, written language makes its appearance; and the child's mastery of that new world is dependent upon two things: the richness and quantity of the language which the child experiences, and the maturation of the child's mental processes which permit him to deal with the logic of language and the abstractness of certain concepts. As the child grows older, written language becomes more important. He consumes written language and produces it at a steadily accelerating rate. And, it seems, language experience is cumulative; it grows by a process of accretion. Several research studies, but particularly that of Hunt⁵, indicate that whatever children do with language at an early age, they continue to do, but as they grow older, they simply add more processes. Thus the structures of oral language are mastered early, or at least are raised to a level of "operational acceptability" early in life. (If the world around a child does not demand exact language, the child will only progress toward mastery so far.) These simpler patterns of language do seem to be used more and more accurately as the child grows older, however. The investigator would suggest that such changes are in response to the greater

⁵ Hunt, op. cit., 1970.

levels of expectancy which the world sets up for the older child or adolescent.

The child's encounters with the patterns of written language, plus the state of maturation of his ability to reason logically and abstractly, plus, once again, the demands which society places upon the child determine his skill with written language. Certain structures of language (for example, three or four levels of self-embedding) boggle the most mature adult. But beyond such structures, the patterns of written language are not terribly difficult in themselves. Some make demands upon logic; others require short-term memory; many require a grasp of shades of dependency (for example, connectives such as "however", "therefore," "consequently," and so forth) and reference. That is to say, the forms of language structures are relatively simple; the manner in which they function is far more difficult.

Some structures of language (that is, some forms of language) can only function in a very special-- and possibly difficult-- way. It is easy to begin thinking too much in terms of the forms which transformations produce instead of realizing that it is the function (for example, the process of negation) that is important. Oral language operates functionally, while written language superimposes, in addition, demands of form and style.

Some children enter early and deeply into the world of written language. Many of them become aware of the special

demands of written discourse, that certain new forms and arrangements of forms are attractive and effective. Consequently, such children adopt these new forms into their own writing style (but rarely or never into their speech). We identify such children as possessing a "more mature style" than their companions who are less print-oriented. Thus an entire continuum of individual differences is introduced into the range of "language ability."

However, at the same time, factors of mental maturity, and social expectation are at work, one limiting the mastery of the functions of certain forms, the other placing ever-greater demands upon the young language user.

As this study indicates, there is a steady developmental trend in the ability of school-age children to use the structures of language. This developmental trend may be attributed to three factors:

1. greater mental maturity,
2. greater exposure to and experience with the structures of language, and
3. social pressures which make greater demands upon older students for accuracy of usage.

The "difficulty" of the structures of language takes the form of a continuum, those at the lower, or easier end of the scale being common to patterns of oral speech, and those at the upper end of the scale being characteristic of written discourse. Further, the arrangement of elements in a written sentence can contribute to a structure's relative difficulty.

What about experiments with the Derivational Theory of Complexity and others showing a graduation of difficulty between various simple transformations and combinations of them? In the experience of the investigator, the question-form is the most familiar of a group including the question, the negative, and the passive. While we are not inclined to speak negatively, the negative is more familiar than the passive. The results of the "complete and delete" test give the following frequency of accurate responses for these three forms:

| | |
|-----------------|-------|
| negative | 74.9% |
| passive | 78.4% |
| yes/no question | 85.8% |
| WH question | 90.5% |

These results do not seem to correspond with the results projected above: the passive should have received a lower frequency of accurate response than the negative. However, a solution is presented by the study done by DeVito⁶, reviewed earlier in section d, section 3 of Chapter II. That study indicated that "cloze" technique makes passives appear easier than actives because of greater redundancy, and an abundance of context clues. Were it not for this factor, the relative positions of the negative and passive -- which are not very far apart in percentage of acceptable response as it is -- might very easily have been reversed.

What might be suggested, then, as a way of determining the relative "ease" with which a particular transformational

⁶Devito, op. cit.

operation can be manipulated would be to prepare a list of relative "familiarity" of such structures. Surely, the structures and their corresponding operation which are common in oral language would be the easiest to use acceptably or correctly. The hardest transformational processes to use would be those from written discourse which involve long clausal embeddings. Combinations of transformations would have an ease of processing which would not be a function of the ease of processing of the individual transform themselves, but which would depend upon the commonness of the combination in question (for example, how often does one run into a negative-passive question?).

This "theory of familiarity" seems very simple on the surface. Nevertheless, how is one to account for the great disparity in frequency of acceptable response between the "it-that" complement structure in the objective position (86.8 percent) and the identical structure used in the subjective position (48.3 percent)? A relative clause in the objective position, a position common in both oral and written discourse, is recognized accurately 86.3 percent of the time; the same structure in the subjective position, a position more common in written language, receives an accurate response figure of 76.6 percent.

Are some of the processes of language which are represented by particular syntactic structures more easily performed by human language users than others? This study clearly seems to answer in the affirmative. Is it possible

to determine the relative "difficulty" with which we process such structures? Again, the somewhat crude methods of this study would seem to answer "yes." Can such difficulty be attributed to a process of node-counting on a transformational diagram, or some theory of derivational history? Everything written in this chapter so far indicates a "no," although there is no reason to doubt that a particularly complicated derivation might result in an unfamiliar structure of language. In that sense, there might conceivably be some correspondence.

Developmental Trends in Syntactic Ability.

The first part of this section has indicated the investigator's interpretation of the processes of development which almost all of the results of the study indicate are present. The easiest transformations, made up largely of structures which are common in oral language, are not completely mastered by grade seven students. The increases in ability with these "easy" processes evidenced in grades nine and eleven are not large, because there simply isn't that much room for improvement. Take, for example, the WH - question transformation number 35. This item was responded to acceptably 86.7 percent of the time by grade seven students, 87 percent by grade nine, and 99.2 percent by grade elevens. Children eight years of age seem to use WH questions accurately. "Where 're yuh goin'?" they ask. "What'cha gonna do, mister?" Yet their response to written items incorporating the WH - question structure does not approach perfection until grade eleven. Can we attribute such a change to increased mental

maturity? Surely the use of WH - questions does not demand abstract thinking. Thus it appears as though we must search elsewhere for this increase in accuracy of performance.

The previous section suggested increased social pressure as an important factor in the increase evident in students' ability to use the structures of language accurately. How else can we account for the developmental trends apparent with so many of the easiest transformational processes studied? The increase in the acceptability of subjects' handling of conjunctions has been mentioned earlier. However, Loban's 1967 study⁷ has indicated that the accurate use of conjunctions and connectives is the mark of the better writer. The investigator would suggest that accurate use of both conjunctions and connectives is not that simple, and that accurate use of such structures/processes makes genuine demands upon reasoning ability. Note further that "conjunction" is near the middle of the list of relative difficulty of transformational processes, and that the developmental pattern which emerges for it is very pronounced across the three grades of the sample.

If social pressure is an important factor working together with increased mental maturity and experience to build language ability within students, where in the student's career might we expect the greatest changes? One result of the present study which was most interesting was the sudden

⁷Loban, op. cit., 1967.

spurt in both syntactic ability and writing ability seemingly put on by the grade elevens. In the case of many transformations, indicated increases in accurate response from grade seven to grade nine were small indeed; but the gap between grade nine and grade eleven is evident.

While the slightly lower average I.Q.'s of the grade nine group might account for some of this discontinuity, a considerable amount remains.

Compared to the insular, inwardly-turned world of junior high school, the grade eleven student's world is far more outwardly directed. The world of work and married life seems much closer. Further, the academic work load and exposure to written material is multiplied several times. The suggestion is that the grade eleven student feels a great deal more pressure to communicate accurately and to use adult patterns of discourse more frequently, particularly in written language. We ask for relatively little written output from junior high school students, and most of what we demand is expository in nature. Further, the reading done from grade nine downwards is mostly narrative, while high school students deal with increasing amounts of non-fiction. All of these factors seem to "snowball" for the student in grades eleven and twelve, with the result that just at the moment the student achieves full development of his mental abilities, he is also immersed in language experiences, and social pressures begin to reach a peak as well. (Students who have been sloppy in their syntax for years realize that

they will soon be writing business letters in which one mistake will make them appear stupid.)

The older the student, the more important the patterns of written language become to him, and the more experience with written language he can bring to bear. More aware of the distinction between oral and written discourse, the older student can even consciously begin to employ the patterns of syntactic structures peculiar to writing, thus widening the gap between himself and the junior high student.

The Relationship of Syntactic Ability and Writing Ability

The writing assignments in this study were collected from three grades: seven, nine, and eleven. Two samples were written by each student in each class, but common topics were used for each of the two groups of papers. The evaluators of those papers were looking for evidence of a writer's ability to use language as an effective means of communication. It is not often that such a cross-grade sample is taken for evaluation. Note that all papers were typed and were identified only by number. The average marks for the three grades were:

Grade seven - 16.07

Grade nine - 17.04

Grade eleven - 21.75

The following chart gives an interesting picture of the two extremes as well. Note that the marks could range from

6 to 30:

TABLE 19
Distribution of High and Low Writing Marks by Grades

| Mark Range | Seven | Grades Nine | Eleven |
|------------|-------|----------------|--------|
| 26 - 30 | - | 3 | 9 |
| 21 - 25 | 9 | 15 | 17 |
| 6 - 10 | 7 | 7 | - |

While it is true that a more mature writer is capable of "doing more" with a given topic, the total absence of any student in the top group for grade seven, and the similar absence of a single grade eleven in the bottom group still indicates that something other than "natural inherent ability," (or lack of it) is operating. Out of 45 grade sevens, not one "natural writer" broke into the top group, while in grade eleven, not a single linguistically-handicapped youngster plunged into the cellar.

A similar table of scores on the "complete and delete" test also shows interesting trends (Table 20). It appears that the upward surge in syntactic ability (which would tend to be ability with the structures of written language because of the weighting system used) is even more evident in grade eleven. But note also that some grade sevens also make the top group, while two grade

elevens remain in the bottom classification. This would seem to indicate that some grade sevens are already sensitive to the patterns of written English, while some grade elevens have never mastered such patterns.

TABLE 20

Distribution of High and Low Syntactic Ability Marks by Grades

| Mark Range | Seven | Grades Nine | Eleven |
|------------|-------|----------------|--------|
| 25 - 30 | 3 | 4 | 12 |
| 21 - 24.9 | 8 | 13 | 13 |
| 7 - 12 | 6 | 3 | 2 |

Comparing these two charts indicates that, despite some syntactic precociousness on the part of some grade sevens they cannot write as well as grade elevens, possibly because they simply lack the linguistic experience. Further, even syntactically inept grade elevens have a quality of maturity in their writing which moves them out of the bottom group. In other words, there is something going on in the way of maturational or experiential influences which a test of syntactic ability cannot measure.

The principal correlation of this study indicates precisely the above hypothesis. The correlation (0.73) is too high for there not to be a considerable part of greater

writing ability reflected in or caused by greater ability with the structures of language and the thought patterns which such structures represent.

C. SUMMARY AND REACTION TO HYPOTHESES

The human child is born with a propensity for language. His first intelligible sounds are the product of imitation. Eventually he becomes aware of his ability to influence and communicate with those around him, and communication becomes the chief motivating factor in his language production, even though he may continue to "experiment" with language in a non-communicative fashion.

The language world of the young child is an oral world. His development within that world is controlled by his rapidly evolving mental maturity, his experience, and a social factor involving his own desire to communicate and his reaction to the expectations and censures of those around him. The syntactic patterns of language he develops are purely those of oral language; and the degree to which he perfects those patterns is directly controlled by the reactions of those with whom he communicates.

Once a child begins to read, he begins to acquire the syntactic patterns of written language, which differ in many respects from the oral patterns with which the child is familiar. The degree to which the child perfects the patterns of written language is controlled by the same factors controlling his acquisition of oral language: his

experience (which must include both reading and producing written language), his mental maturity, and the same social factors mentioned above.

Only by means of rich and expanding experiences with written language can the child acquire mastery of that medium. His motivation to read must be from within himself-- a desire for knowledge, entertainment, or whatever. However, his motivation for writing must be a desire to communicate, to influence those about him in some meaningful way.

The young writer will use the most familiar pattern of language he knows to express an idea, until he is aware of alternative patterns, becomes familiar with them, and can see some functional or stylistic reason for using them. There are no "difficult" patterns of language for the young writer, only those with which he is not familiar, or whose function or utility he does not understand.

Students differ in their ability with language of all kinds because they lack mental ability, or linguistic experience, or the personal or social factors of motivation and control.

A young writer's syntactic ability reflects his experience with both oral and written language. However, there is no one-to-one correspondence between ability as a writer and ability with the structures of language. Even though a young writer may be aware of the possibility of using a different structure of language to express an idea or a group of ideas, he may not choose to use that structure,

or he may not see in what way that alternative is more desirable. To be "competent" with the structures of language does not necessarily mean that the entire range of such structures will appear in performance.

D. LIMITATIONS OF APPLICABILITY OF FINDINGS

The sample of this study was drawn in an arbitrary fashion from schools assigned to the investigator. Consequently, it could not be called a true random sample. Variations in socio-economic backgrounds and the range of mental abilities represented were wide indeed. The effect of these factors would not be great upon the correlation of individual student's syntactic ability mark and writing mark, but they may very well have had some influence upon the developmental aspect of the study.

The range of syntactic structures which were investigated was definitely limited in order that the test instrument might not become cumbersome. There are undoubtedly other such structures which should be investigated in a similar fashion. Further, it is entirely probable that many other transformational structures were present in the test instrument, and had their influence on the frequency of acceptable answers obtained. Another factor which the study reveals to be of importance in determining the relative "difficulty" of the test items is their position within their "matrix" sentence. Except for certain of the complement structures, no effort was made to control for

this potentially important factor.

One potential source of inaccuracy of findings, was the incorporation of more than one "cloze" blank in a single sentence. The investigator felt that on several occasions, mistakes on one item led to mistakes on others. A better approach might be to limit test items to one to a sentence.

The reported "scale of difficulty" of the structures incorporated in the instrument of this test is thus peculiar to the test itself and the passages within which the test items were incorporated.

A particular weakness of the test instrument was the approach taken to deletion structures, an approach which seemed unable to discriminate either the relative difficulty of structures or any clear developmental trends in the use of such structures.

The topics chosen for the writing assignments were capable of stimulating a response from all the subjects of the test population. However, no effective way of controlling the manner in which the subjects chose the content of their responses, nor of controlling the total length of their replies was conceived. Both of these factors may have contributed to the final marks assigned to subjects.

The writing sample for each subject was limited to from 200 to approximately 500 words, and there was a time limit for writing as well. The adequacy of a sample of this length written under a time limit is potentially questionable.

E. SUGGESTIONS FOR FURTHER RESEARCH

This entire study could very well be replicated using an instrument which concentrated upon a narrower range of structures of language — possibly embedding structures — in order to reduce the influence of non-discriminating items on total overall score.

The test instrument of this study, and "cloze procedure" in general, are worth testing further in order to examine the structures of language. No other approach seems capable of utilizing context as effectively. However, further studies should take the positions of structures within sentences into consideration, and should also limit test items to one per sentence in order to avoid interference of one item with another.

A similar study should be done on both younger and older subjects in order to establish the validity of the hypothesis put forward in the latter part of this investigation that language structures common in speech are far more familiar to younger writers, but are eventually mastered by the mature writer, while development of the structures peculiar to written language lags in young writers, and undergoes steady expansion as the writer matures.

A similar study incorporating a much more adequate sample of writing from all subjects should be conducted, with the investigator then examining the written structures of the writing sample to see if, in fact, the subjects did

use the structures which they responded to acceptably on the "Cloze" type test.

F. IMPLICATIONS OF THE STUDY

Clearly, students who have mastered the syntactic structures of language, and particularly those of written language, have an excellent chance of being judged better writers. Should we consequently begin a program designed to give students practice with the "embedding" structures which appear in the writing of better writers? Such a practice would expose students to a different concept with respect to the processes of modification, a concept which is in itself old, but which modern transformational theory makes more amenable to teaching. No doubt, an ingenious and resourceful teacher could devise such an approach and could probably get results. With luck, some of the students might even grasp the effects of such processes of modification, and be able to discriminate as to appropriate places to insert such forms.

The entire process of building genuine language ability must, however, be begun much earlier, and much more fundamentally. Superimposing an ability to combine sentences upon the "skills" of a youngster who writes sentence fragments seems fruitless at best.

Discussion of Related Concepts and Objectives

The first and most important step in developing youngsters skilled in language is to raise our expectations, as parents and teachers, with respect to their oral language

skills. So long as we maintain our low expectations of skill, children will function linguistically at that low level, the level of bare communication. While it is true that children and teenagers as well as adults can "shift gears," and fit their mode of discourse to the situation at hand, still, we should not then wonder at the difficulty with which they become "skilled writers." We should not demand a level of oral discourse which is totally inappropriate, but we must not accept the lowest common denominator of communication either.

The teacher particularly is in a position to demand and get a little more. Further, the teacher herself must be a good model. Indeed, "speech" must have a vital place in the elementary school curriculum, on a formal, as well as an informal basis. The child should be made aware of different "levels" of discourse, and their appropriateness in various situations.

The key to developing truly skilled users of written language lies in the attitude which is fostered in the child toward writing and the printed word. If a child reads only to write examinations, or to meet "requirements" he will not be a happy reader. If books mean only information and never enjoyment, the child will turn to books only "when he has to," to get an assignment out of the way. Not only must every effort be made to make reading pleasurable to the child, but the development of the child's appreciation of books should become central in the process of education during the

early stages. Until some other form of getting information and vicarious experience replaces books, they will remain the principal source of knowledge in all fields.

No matter how much a child reads, however, his skill with language will be developed only when he can write skilfully and with enjoyment. And he will write joyfully only if he has some reason other than "because he has to" for writing. Children speak because they want to communicate, and because they want to influence those about them. They want to get something, or tell somebody about something, or possibly just make a noise. They will never write anything with pleasure, nor will they begin to develop their potential as writers, until the entire process of written communication becomes worthwhile in and for itself, and for what it can do for them.

The young writer's skill with all of the structures of language can only be attained if he is familiar with those structures, and can use them effectively and with the knowledge that he is using them appropriately. Such knowledge can only come from experience guided by teachers who are themselves knowledgeable about the structures of language and able to recognize effective language use. The most important source of the child's experience must come from his own encounters with the printed word. But the awareness of logic, structure, and style must come from teachers who know language and rhetoric, and who can establish in the mind of the child the desirability of skill

with language.

Where the child will make such knowledge and skill operable is in his own writing. He must write, and write, and be self-critical, and then write some more. His writing should not be so much "marked" as evaluated against his previous efforts, and occasionally admired or even informally published for his classmates, friends, and parents to admire as well.

Only when skilled writing becomes pleasurable and something worth striving for will the young writer reach his potential. The structures of language will then become his functional tools, the tools of a craftsman.

G. CONCLUDING STATEMENT

This study began as an investigation of some of the structures of language. Part-way through its gestation period, it was extended to examine student writing as well. At its conclusion, it has become a recognition of the "human" quality of language. The structures of language have been recognized as being only the functional tools of the writer's ingenuity, the products of his mental capacity and his linguistic experience. How well the students that participated in this study wrote was not determined by the topics they were given, nor particularly by how they happened to feel that day. Each participating student was revealing, by means of what he put down and how he put it down, what had happened to him as a user of language for the past

fifteen or seventeen years.

Some of the papers were masterful; some were even full of joy; some were lackluster-- weak attempts at satisfying "another teacher." And some were truly pitiful, the products of language cripples cut off from the world of ideas which is written language.

I apologize to all of them, for I did little to make their experience with the world of language any more pleasant. However, perhaps this study can help to shed some small light upon the tasks which confront the English teacher, and contribute toward our knowledge of the processes of language.

BIBLIOGRAPHY

BIBLIOGRAPHY

A. Books

- Allen, Harold B. (ed.) Readings In Applied Linguistics.
Second Edition. New York: Appleton-Century-Crofts, 1964.
- Bach, E. An Introduction to Transformational Grammars.
New York: Holt, Rinehart, and Winston, Inc., 1964.
- Bach, Emmon and Robert T. Harms (eds.) Universals in
Linguistic Theory. New York: Holt, Rinehart, and
Winston, Inc., 1968.
- Bateman, Donald and Frank Zidonis. The Effect of a Study
of Transformational Grammar on the Writing of Ninth
and Tenth Graders. Champaign, Illinois: National
Council of Teachers of English, 1966.
- Bateman, D. R. Speculations Concerning Symbolism, the
Communications Core, and Language. Columbus: Center
for School Experimentation, the Ohio State University,
1959.
- Bloomfield, Leonard. Language. New York: Holt, Rinehart
and Winston, 1933.
- Braddock, Richard et. al. Research in Written Composition.
Champaign, Illinois: National Council of Teachers of
English, 1963.
- Chomsky, Noam. Cartesian Linguistics. New York: Harper
and Row, 1966.
- _____. Aspects of the Theory of Syntax. Cambridge,
Mass.: M. I. T. Press, 1965.
- _____. Current Issues in Linguistic Theory. The Hague:
Mouton and Company, 1964.
- _____. Syntactic Structures. The Hague: Mouton and
Company, 1957.
- Chomsky, Noam and Morris Halle. The Sound Pattern of
English. New York: Harper and Row, 1968.
- Christensen, Francis. Notes Toward A New Rhetoric. New
York: Harper and Row, 1967.

- Davis, Edith A. The Development of Linguistic Skill in Twins, Singleton with Siblings, and Only Children of Age Five to Ten Years. Institute of Child Welfare Monograph Series, No. 14. Minneapolis: University of Minnesota Press, 1937.
- Diedrich, P. G. Factors in the Judgment of Writing Ability. E. T. S. Research Bulletin 61-65. Princeton, New Jersey: E. T. S., 1961.
- Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Company, 1966.
- Flesch, R. F. Marks of a Readable Style: A Study in Adult Education. New York: Bureau of Publications, Teachers' College, Columbia University, 1943.
- Fodor, Jerry A. and Jerrold J. Katz. The Structure of Language. Englewood-Cliffs, New Jersey: Prentice-Hall, Inc., 1964.
- Fries, Charles C. The Structure of English. New York: Harcourt, Brace and World, 1952.
- Gage, N. L. (ed.) Handbook of Research on Teaching. New York: Rand, McNally, 1963.
- Gaskin, James R. and Jack Suberman. A Language Reader for Writers. Englewood Cliffs, New Jersey: Prentice-Hall, inc., 1966.
- Gleason, H. A. Jr. Linguistics and English Grammar. New York: Rinehart and Winston, 1965.
- Gunning, R. The Technique of Clear Writing. New York: McGraw-Hill Book Company, 1952.
- Harris, Zellig S. Methods in Structural Linguistics. Chicago: The University of Chicago Press, 1951.
- Hunt, Kellogg. Syntactic Maturity in School Children and Adults. Monographs for the Society For Research in Child Development, XXXV No. 1, Chicago: University of Chicago Press, 1970.
- Grammatical Structures Written at Three Grade Levels. Champaign, Illinois: National Council of Teachers of English, 1965.
- Jacobs, Roderick A. and Peter S. Rosenbaum. English Transformational Grammar. Waltham, Massachusetts: Blaisdell Publishing Company, 1968.

- Jespersen, Jens Otto Harry. A Modern English Grammar on Historical Principles. Copenhagen: Ejnar Munksgaard, 1909-1949.
- _____. Essentials of English Grammar. London: George Allen and Unwin, Limited., 1933.
- Katz, J. J. and P. M. Postal. An Integrated Theory of Linguistic Descriptions. Research Monograph No. 26. Cambridge, Mass.: M. I. T. Press, 1964.
- Kruisinga, Etsko. A Handbook of Present-Day English. Fourth Ed., Utrecht: Kemink en Zoon, 1925
- Labov, William. The Study of Nonstandard English, Champaign, Illinois: National Council of Teachers of English, 1970.
- Lamb, Sydney M. Outline of Stratificational Grammar. Washington, D. C. Georgetown University Press, 1966.
- Lees, Robert B. A Grammar of English Nominalizations. Bloomington, Indiana: Research Centre in Anthropology, Folklore, and Linguistics, 1960.
- Lefevre, Carl A. Linguistics, English, and the Language Arts. Boston: Allyn, Bacon Inc., 1970.
- Loban, Walter. The Language of Elementary School Children. National Council of Teachers of English Research Report No. 1. Champaign, Illinois: National Council of Teachers of English, 1963.
- Long, Ralph B. The Sentence and Its Parts. Chicago: The University of Chicago Press, 1961.
- Lowth, Robert. A Short Introduction to English Grammar. (London, 1762). Reprinted: Menston, England: Scholar Press, 1967.
- Lyman, Rollo L. Summary of Investigations Relating to Grammar, Language and Composition. Chicago: University of Chicago Press, 1929.
- Lyons, J. and R. J. Wales, (eds.) Psycholinguistic Papers. Edinburgh: The University of Edinburgh, 1966.
- McCarthy, Dorothea A. Language Development of the Preschool Child. Institute for Child Welfare Monograph Series, No. 4. Minneapolis: University of Minnesota Press, 1930.

- McNemar, Quinn. Psychological Statistics. Third Edition. New York: John Wiley and Sons, 1962.
- Nida, Eugene Albert. A Synopsis of English Syntax. Norman, Oklahoma: Summer Institute of Linguistics, 1960.
- Nist, John. A Structural History of English. London: St. Martin's Press, 1966.
- O'Donnell, Roy C., William J. Griffen, and Raymond C. Norris. Syntax of Kindergarten and Elementary School Children: A Transformational Analysis. National Council of Teachers of English Research Report No. 8. Champaign, Illinois: National Council of Teachers of English, 1968.
- Potter, Simeon. Modern Linguistics. New York: W. W. Norton and Company, Inc., 1964.
- Poutsma, Hendrik. A Grammar of Late Modern English. Second Edition. Groningen: P. Noordhoff, 1914-1929.
- Prideaux, Gary. The Syntax of Japanese Honorifics. The Hague: Mouton and Company, 1969.
- Rosenbaum, Peter S. The Grammar of English Predicate Complement Constructions. Cambridge, Mass.: M. I. T. Press, 1967.
- Saint-Jacques, Bernard. Analyse Structurale de la Syntaxe du Japonais Moderne. Paris, 1966.
- Sapir, Edward. Language: An Introduction to the Study of Speech. New York: Harcourt, Brace, 1921.
- Saporta, S. (ed.) Psycholinguistics: A Book of Readings. New York: Holt, Rinehart, and Winston, Inc., 1962.
- Terwilliger, Robert F. Meaning and Mind. New York: Oxford University Press, 1968.
- Thorndike, E. L. The Teacher's Word Book. New York: Teachers' College. Columbia University, 1921.
- Thomas, Owen. Transformational Grammar and the Teacher of English. New York: Holt, Rinehart, and Winston, Inc., 1965.
- Watts, A. F. The Language and Mental Development of Children. Boston: D. C. Heath and Company, 1948.
- Zandvoort, Willem Reinard. A Handbook of English Grammar. London: Longmans, 1957.

B. Articles

- Allen, Harold B. "From Prairies to Mountains," College English. XXVI (January, 1965), 260-266.
- Ammon, P. R. "The Perception of Grammatical Relations in Sentences: A Methodological Exploration." Journal of Verbal Learning and Verbal Behavior. V (1968), 869-875.
- _____. "Some Research on the Comprehension of Sentences." Child Study, XXX (Winter, 1968-69), 14-18.
- Anderson, C. C. "The New STEP Essay Test as a Measure of Comprehension Ability." Educational and Psychological Measurement, XX (Spring, 1960), 95-102.
- Anglin, J. M. and G. A. Miller. "The Role of Phrase Structure in the Recall of Meaningful Verbal Material." Psychonomic Science, 1968, X, 343-344.
- Bear, Meta. "Children's Growth in the Use of Written Language." Elementary English Review, XVI (December, 1939), 312-319.
- Bever, T. G., J. R. Lackner, and L. Kirk. "The Underlying Structures of Sentences are the Primary Units of Immediate Speech Processing." Perception and Psycholinguistics, 1969 (4), 225-234.
- Bickley, A. C., E. E. Weaver, and F. G. Ford. "Information Removed From Multiple-Choice Item Responses by Selected Grammatical Categories." Psychological Reports, XXIII (1968), 613-614.
- Bormuth, J. R. "Empirical Determination of the Instructional Reading Level." International Reading Association Conference Proceedings, XIII (1969), Part I, 716-721.
- _____. "Cloze Readability Procedure." CSEIP Occasional Report, No. 1, University of California, Los Angeles, October, 1967.
- _____. "Readability: A New Approach." Reading Research Quarterly, I (Spring, 1966), 79-132.
- Bradley, Michael. "Effects on Reading Tests of Deletions of Selected Grammatical Categories." Reading: Process and Pedagogy, Nineteenth Yearbook of the National Reading Conference, Vol. 1, Process, 87-93.

- Carroll, John B. "Language Development in Children," in Encyclopedia of Educational Research, ed. Chester W. Harris, New York: Macmillan, 1960, 744-752.
- _____. "Linguistics and the Psychology of Language." Review of Educational Research, 34: 119-126 (1964).
- Cast, B. M. D. "The Efficiency of Different Methods of Marking English Composition." British Journal of Educational Psychology, IX (November, 1939), 257-269, and X (February, 1940), 49-60.
- Chomsky, Noam. "Topics in the Theory of Generative Grammar." in T. Sebeok, ed., Current Trends in Linguistics, III, Linguistic Theory. Bloomington: Indiana University Press, 1966.
- _____. "The Current Scene in Linguistics." College English, XXVII (May, 1966), 587-595.
- _____. "A Transformational Approach to Syntax," In A. A. Hill, (ed.), Proceedings of the 1958 Conference on Problems of Linguistic Analysis in English, Austin, Texas: 1962, 124-148.
- _____. "Some Methodological Remarks on Generative Grammar." Word XVII (1961), 219-239.
- _____. "A Review of B. F. Skinner's Verbal Behavior." (New York: Appleton-Century-Crofts, Inc., 1957) in Language, XXXVI, No. 1 (1959), 26-58.
- Christensen, Francis. "The Problem of Defining a Mature Style." English Journal, LVII (April, 1968), 572-579.
- _____. "A Generative Rhetoric of the Sentence." College Composition and Communication, XVI (1963), 151-161.
- Coleman, E. B. "Learning of Prose Written in Four Grammatical Transformations." Journal of American Psychology, XLIX (1965), 332-341.
- Coleman, E. B. and J. P. Blumfield. "Cloze Scores of Nominalizations and their Grammatical Transformations." Psychological Reports, XIII (1963), 651-654.
- Compton, Arthur J. "Aural Presentation of Different Syntactic Structures and Lengths." Language and Speech, X (April, 1967), 81-87.

- Cronbach, Lee J., Nageswari Rajaratnam, and Goldine C. Gleser. "Theory of Generalizability: A Liberalization of Reliability Theory." British Journal of Statistical Psychology, XVI(21, November, 1963), 137-163.
- Culhane, J. W. "Cloze Procedures and Comprehension." The Reading Teacher, XXIII(February, 1970), 410-413.
- Dale, E. and J. S. Chall, "A Formula for Predicting Readability." Educational Research Bulletin, XXVII (January, 1948), 11-20.
- Dale, E. and R. W. Tyler, "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability." Library Quarterly, IV(July, 1934), 384-412.
- Davidson, Robert E. "Transitional Errors and Deep Structure Differences." Psychonomic Science, XVI(1969), 293-294.
- DeVito, Joseph A. "Some Psycholinguistic Aspects of Active and Passive Sentences," Quarterly Journal of Speech, LV(December, 1969), 401-406.
- Diederich, Paul. "The Measurement of Skill in Writing," School Review, LIV(December, 1946), 586-587.
- Dolch, E. W. "Vocabulary Burden," Journal of Educational Research, XVII (March, 1928), 170-183.
- Dusel, William J. "Some Semantic Implications of Theme Correction." English Journal, XXXIV(1955), 390-397.
- Epstein, William. "Some Conditions of the Influence of Syntactical Structure on Learning:Grammatical Transformations Learning Instructions, and 'Chunking'." Journal of Verbal Learning and Verbal Behavior, VI(1967), 415-419.
- _____. "Temporal Schemata in Syntactically Structured Material." Journal of General Psychology, LXVIII(1963), 157-164.
- _____. "A Further Study of the Influence of Syntactical Structure on Learning." American Journal of Psychology, LXXV(1962), 121-126.
- _____. "The Influence of Syntactical Structure on Learning." American Journal of Psychology, LXXIV(1961), 81.
- Ervin, S. M. "Changes With Age in Verbal Determinants of Word Association." American Journal of Psychology, LXXIV(1961), 361-372.

- Ervin, Susan M. and W. R. Miller. "Language Development," in Child Psychology, Sixty-Second Yearbook Part 1, National Society for the Study of Education, ed. Harold W. Stevenson, et al. Chicago: University of Chicago Press, 1963, 108-143.
- Farr, James N., James J. Jenkins, and Donald G. Paterson, "Simplification of the Flesch Reading Ease Formula." Journal of Applied Psychology, XXXV (October, 1951), 333-337.
- Fillmore, Charles J. "The Case for Case," in Emmon Bach and Robert T. Harms, (eds.) Universals in Linguistic Theory. New York: Holt, Rinehart and Winston, Inc., 1968, 1-88.
- Finlayson, D. S. "The Reliability of the Marking of Essays." British Journal of Educational Psychology, XXI (1951), 216-234.
- Flesch, R. F. "A New Readability Yardstick." Journal of Applied Psychology, XXXII (June, 1948), 221-233.
- Fodor, J. A. and M. Garrett. "Some Reflections on Competence and Performance." in J. Lyons and R. J. Wales (eds.) Psycholinguistic Papers. Edinburgh: University of Edinburgh, 1966.
- Fodor, J. A. and T. G. Bever. "The Psychological Reality of Linguistic Segments." Journal of Verbal Learning and Verbal Behavior, IV (1965), 414-420.
- Fodor, J. A. and M. Garrett, "Some Syntactic Determinants of Sentential Complexity." Perception and Psychophysics, II (July, 1967), 289-296.
- Forster, Kenneth I. "The Effect of Syntactic Structure on Nonordered Recall," Journal of Verbal Learning and Verbal Behavior, V (1966), 292-297.
- Frogner, Ellen. "Problems of Sentence Structure in Pupils' Themes," English Journal, XXII (November, 1933), 742-749.
- Garrett, M., T. A. Bever, and J. A. Fodor. "The Active Use of Grammar in Speech Perception," Perception and Psycholinguistics, I (1966), 30-32.
- Glanzer, M. "Grammatical Category: A Rote Learning and Word Association Analysis," Journal of Verbal Learning and Verbal Behavior, I (1962), 31-41.

- Gleason, J. A. Jr., "The Organization of Language; A Stratificational View." Report of the Fifteenth Annual (First International) Round Table Meeting on Linguistic and Language Studies, Ed. C.I.J.M. Stuart. Washington, D.C.: Georgetown University Press, 1964.
- Gleser, G. C., L. J. Cronbach, and N. Rajaratnam, "Generalizability of Stratified Parallel Tests," Psychometrika, XXX, No. 1 (March, 1963), 39-56.
- Gough, Phillip B. "The Verification of Sentences: The Effect of Delay of Evidence and Sentence Length." Journal of Verbal Learning and Verbal Behavior, V(1966), 492-496.
- Graf, Richard and Jane W. Torrey, "Perception of Phrase Structure in Written Language." Proceedings of the Seventy-Fourth Annual Convention of the American Psychological Association, Inc., 1966, 83-84.
- Harrell, Lester E. Jr. "A Comparison of Oral and Written Language in School-Age Children." Monographs of the Society for Research in Child Development, XXII(1957).
- Heider, F. K. and G. M. Heider, "A Comparison of Sentence Structure of Deaf and Hearing Children." Psychological Monographs, LII(1940), 42-103.
- Herriot, Peter. "The Comprehension of Syntax." Child Development, 39 (March, 1968), 273-282.
- Hocker, Mary Elsa. "Reading Materials for Children Based on Their Language Patterns of Syntax, Vocabulary, and Interests," (Unpublished Masters thesis, University of Arizona, 1961).
- Ivic, Milka. "Non-admissable Determiners in Slavic Languages." Proceedings of the Ninth International Congress of Linguists. The Hague: 1966, 476-479.
- Kirby, C. L. "Using the Cloze Procedure as a Testing Technique," International Reading Association Conference Proceedings, XIII(1970), Part 4, 68-77.
- LaBrant, Lou. "A Study of Certain Language Developments of Children in Grades Four to Twelve, Inclusive." Genetic Psychology Monographs, XIV (November, 1933), 387-497.
- Lakoff, George. "On Generative Semantics." (Unpublished manuscript).

- Lively, B. A. and S. L. Pressey. "A Method of Measuring the 'Vocabulary Burden' of Textbooks." Educational Administration and Supervision, IX (October, 1923), 389-398.
- Lorge, I. "The Lorge and Glesch Readability Formulae: A Correction." School and Society, LVII (February 21, 1948), 141-142.
- _____. "Predicting Reading Difficulty of Selection for Children." Elementary English Review, XVI (October, 1939), 229-233.
- Marks, Lawrence E. and George A. Miller. "The Role of Semantic and Syntactic Constraints in the Memorization of English Sentences." Journal of Verbal Learning and Verbal Behavior. III (1964), 1-5.
- McCarthy, Dorothea. "Language development in Children," in Manual of Child Psychology, Second edition., ed. Leonard Carmichael, New York: John Wiley and Sons, 1954, 492-630.
- McCuaig, Roger A. "How Not to Analyze the Syntax of Children: A Critique and a Proposal." Elementary English, XLIII (May, 1970), 612-618.
- Mehler, Jacques. "Some Effects of Grammatical Transformations on the Recall of English Sentences." Journal of Verbal Learning and Verbal Behavior, II (1963), 346-351.
- Mehler, Jacques and Peter Carey. "The Role of Surface Structure and Base Structure in the Perception of Sentences." Journal of Verbal Learning and Verbal Behavior, VI (1967), 335-338.
- Menyuk, Paula. "Syntactic Rules Used by Children from Preschool Through First Grade." Child Development, 35:2 (June, 1964), 533-546.
- Miller, G. R. and E. B. Coleman. "A Set of 36 Prose Passages Calibrated for Complexity." Journal of Verbal Learning and Verbal Behavior, VI (1967), 851-854.
- Miller, George. "The Psycholinguists." Encounter, Vol. 23, No. 1 (July, 1964), 29-37.
- Miller, G. A. and K. McKean. "A Chronometric Study of Some Relations Between Sentences." Quarterly Journal of Experimental Psychology, XVI (1964), 297-308.
- Miller, George A. and Stephen Isard. "Some Perceptual Consequences of Linguistic Rules." Journal of Verbal Learning and Verbal Behavior, II (1963), 217-228.

- Miller, George A. "Some Psychological Studies of Grammar." American Psychologist, XVII(1962), 748-762.
- Miller, G. A., K. McKean, and D. Slobin. "The Exploration of Transformations by Sentence Matching," in G. A. Miller, "Some Psychological Studies of Grammar." American Psychologist, XVII(1962), 748-762.
- Miller, George A., George Heise, and William Lichten, "The Intelligibility of Speech and Function of the Context of the Test Materials." Journal of Experimental Psychology, XLI(1951), 329-333.
- Miller, G. A. "The Magic Number of Seven, Plus or Minus Two." Psychological Review, LXIII(1956), 81-97.
- Morris, Val A., Fred C. Rakine, and Arthur S. Reber, "Sentence Comprehension, Grammatical Transformations, and Response Availability." Journal of Verbal Learning and Verbal Behavior, VI(1968), 1113-1115.
- Mosberg, L. T. C. Potter, and R. K. Cornell. "The Relation Between Cloze and Multiple Choice Test Scores as a Function of Relative Paragraph Difficulty and Grade Level." Southwest Regional Laboratory Technical Report, No. 6, 1968.
- Nisbet, J. D. "English Composition in the Secondary School Selection." British Journal of Educational Psychology, XXXV(1955), 51-54.
- Noyes, Edward S. "Essay and Objective Tests in English." College English Review, No. 49(Winter, 1963), 10.
- Osgood, C. J. E. and T. A. Sebook, (eds.). "Psycholinguistics: A Survey of Theory and Research Problems." International Journal of American Linguistics, 1954, 20:(4), Memoir II.
- Potter, Robert R. "Sentence Structure and Prose Quality." Research in the Teaching of English, I(1967), 17-28.
- Perfetti, Charles A. and Doba Goodman. "Memory for Sentences and Noun Phrases of Extreme Depth." Quarterly Journal of Experimental Psychology, XXIII(1971), 22-33.
- Prideaux Gary D. "On The Notion 'Linguistically Significant Generalization'." Lingua, 26 (1971), 337-347.

- Rankin, E. F. "The Cloze Procedure - Its Validity and Utility." The Eighth Yearbook of the National Reading Conference, Milwaukee: The National Reading Conference, Inc., 1959.
- Rohrman, Nicholas L. "The Role of Syntactic Structure in the Recall of English Nominalizations." Journal of Verbal Learning and Verbal Behavior, VII(1968), 904-912.
- Ruddell, Robert B. "The Effect of Oral and Written Patterns of Language Structure on Reading Comprehension." The Reading Teacher, XVIII(1965), 270-275.
- Salinger, Kurt and Carol Eckerman. "Grammar and the Recall of Chains of Verbal Responses." Journal of Verbal Learning and Verbal Behavior, VI(1967), 232-239.
- Savin, Harris B. and Ellen Perchonock. "Grammatical Structure and the Immediate Recall of English Sentences." Journal of Verbal Learning and Verbal Behavior, IV(1965), 348-353.
- Schleslinger, I. M. "The Influence of Sentence Structure on the Reading Process." Technical Report, Number 24, United States Office of Naval Research, Information Systems Branch, Jerusalem, Israel, October, 1966.
- Seegers, J. C. "Form of Discourse and Sentence Structure." Elementary English Review, X(March), 1933, 51-54.
- Skyler, G. "Comparison of the Cloze Procedure and the Short Answer Achievement Test in Determining Subject Matter Comprehension." Graduate Research in Education, V(Fall, 1969), 29-45.
- Slobin, Dan I. "Grammatical Transformations and Sentence Completion in Childhood and Adulthood." Journal of Verbal Learning and Verbal Behavior, V(1969), 219-222.
- Spache, G. "A New Readability Formula for Primary-Grade Reading Material." Elementary School Journal, LIII (March, 1953), 410-413.
- Suci, George J., Paul Ammon, and Peter Gamlin. "The Validity of the Probe Latency Technique for Assessing Structure in Language." Language and Speech, X(1967), 69-80.
- Taylor, Wilson L. "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Attitude." Journal of Applied Psychology, XLI(February, 1957), 19-26.
- _____. "Recent Developments in the Use of 'Cloze Procedures'." Journalism Quarterly, XXXIII(1956), 34-46.

- _____. "Cloze Procedure: A New Tool for Measuring Readability." Journalism Quarterly, XXX(Fall, 1953).
- Vygotsky, L. S. "Thought and Speech." Chapter VII of Language and Thought, tr. Helen Kogan, Eugenia Hanfmann, and Jacob Kasanin, in Psycholinguistics: a Book of Readings, ed. Sol Saporta, New York: Holt, Rinehart, and Winston, 1961, 508-527.
- Washburne, C., and M. V. Morphett. "Grade Placement of Children's Books." Elementary School Journal, XXXVIII(January, 1938), 355-364.
- Watt, W. C. "On Two Hypotheses Concerning Psycholinguistics." J. R. Hayes (ed.) Cognition and the Development of Language. New York: John Wiley and Sons, Inc., 1970.
- Wiseman, Stephen and Jack Wrigley. "Essay Reliability: The Effect of Choice of Essay Title." Educational and Psychological Measurement, XVIII(Spring, 1958), 129-138.
- Wiseman, Stephen. "The Marking of English Composition in Grammar School Selection." British Journal of Educational Psychology, XIX(November, 1949), 200-209.
- Wright, Patricia. "Transformations and the Understanding of Sentences." Language and Speech, III(1969), 156-166.
- Yngve, V. "A Model and a Hypothesis for Language Structure." Proceedings of the American Philosophical Society, 140(1960), 444-466.
- C. Unpublished Materials
- Buxton, Earl W. "An Experiment to Test the Effects of Writing Frequency and Guided Practice upon Students' Skill in Writing Expressions." (Unpublished Doctoral dissertation, Stanford University, 1958).
- Fagan, William T. "An Investigation Into the Relationship Between Reading Difficulty and the Number and Types of Sentence Transformations." (Unpublished Doctoral dissertation, University of Alberta, 1969).
- Fletcher, J. E. "A Study of the Relationships Between Ability to Use Context as an Aid in Reading and Other Verbal Abilities." (Unpublished Doctoral dissertation, University of Washington, 1959).

- Gajadharsingh, Joel Lancelot. "A Study of the Effects of Instruction in the Rhetoric of the Sentence on the Written Composition of Junior High School Students." (Unpublished Doctoral dissertation, University of Alberta, 1970).
- Hunt, Kellogg W. "Sentence Structure Used by Superior Students in Grades Four, and Twelve, and By Superior Adults." Report to the U. S. Office of Education, Cooperative Research Project No. 5-0313, 1966.
- _____. "Differences in Grammatical Structures Written at Three Grade Levels, the Structures to be Analyzed by Transformational Methods." Report to the U. S. Office of Education, Cooperative Research Project No. 1998. Tallahassee, Florida: 1964.
- Jenkinson, Marion D. "Selected Processes and Difficulties of Reading Comprehension." (Unpublished Doctoral dissertation, University of Chicago, 1957).
- Kincaid, Gerald L. "Some Factors Affecting Variations in the Quality of Students' Writing." (Unpublished Ed. D. dissertation, Michigan State University, 1953).
- Loban, Walter D. "Language Ability: Grades Ten, Eleven, and Twelve." Report to the U. S. Office of Education, Cooperative Research Project No. 2387. Berkeley: 1967.
- _____. "Language Ability, Grades Seven, Eight, and Nine." Report to the U. S. Office of Education, Cooperative Research Project No. 1131. Berkeley: 1964.
- _____. "Language Ability in the Middle Grades of Elementary School." Report to the U. S. Office of Education, Contract 7287. Berkeley: 1961.
- McMahon, E. "Grammatical Analysis as Part of Understanding." (Unpublished Doctoral dissertation, Harvard University, 1963).
- Mellon, John C. "Transformational Sentence-Combining: A Method For Enhancing the Development of Syntactic Fluency in English Composition." Final Report, Cooperative Research Project No. 5-8418, U. S. Department of Health, Education, and Welfare, Office of Education, 1967.
- Prideaux, Gary. "Sample English Grammars." (Unpublished material, Linguistics 404, University of Alberta, 1970).

- Rawson, Hildred I. "A Study of the Relationships and Development of Reading and Cognition." (Unpublished Doctoral dissertation, University of Alberta, 1969).
- Riling, Mildred E. "Oral and Written Language of Children in Grades 4 and 6 Compared with the Language of Their Textbooks." Report to the U. S. Office of Education, Cooperative Research Project No. 2410. Durang, Oklahoma: 1965.
- Robertson, J. E. "An Investigation of Pupil Understanding of Connectives in Reading." (Unpublished Doctoral dissertation, University of Alberta, 1966).
- Sam, N. H., and E. S. Stine. "Structural Analysis of the Written Composition of Intermediate Grade Children." Report to the U. S. Office of Education, Cooperative Research Project No. S-057. Bethlehem, Pennsylvania: 1965.

APPENDICES

APPENDIX A

Raw Scores for Writing Assignments, and "Complete" Test, Plus Age, Sex, Intelligence Scores, and Self-Rating in English Scores: By Grades

NOTE: "1" indicates male, "2" indicates female.

Self-Rating: "1" indicates "poor"; 5, "superior."

1. GRADE SEVEN

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|-------|--------------|-----|-------------|
| 129 | 21.636 | 24 | 123 | 150 | 2 | 4 |
| 55 | 18.496 | 15 | 106 | 158 | 2 | 4 |
| 27 | 14.329 | 14 | 95 | 160 | 2 | 3 |
| 110 | 21.399 | 12 | 115 | 157 | 1 | 4 |
| 140 | 19.779 | 10 | 145 | 149 | 1 | 4 |
| 13 | 13.515 | 16 | 108 | 162 | 2 | 2 |
| 5 | 20.810 | 23 | 116 | 150 | 1 | 3 |
| 169 | 22.777 | 16 | 125 | 148 | 1 | 4 |
| 151 | 23.820 | 19 | 114 | 154 | 1 | 3 |
| 35 | 16.002 | 15 | 106 | 151 | 1 | 3 |
| 45 | 26.727 | 24 | 147 | 149 | 2 | 4 |
| 161 | 26.100 | 20 | 146 | 154 | 2 | 4 |
| 20 | 16.072 | 18 | 119 | 158 | 1 | 3 |
| 124 | 20.612 | 20 | 140 | 139 | 1 | 4 |

Grade Seven (Continued).

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|-------|--------------|-----|-------------|
| 32 | 16.939 | 19 | 97 | 148 | 1 | 3 |
| 9 | 9.135 | 12 | 85 | 164 | 1 | 1 |
| 63 | 15.798 | 20 | 122 | 156 | 2 | 3 |
| 86 | 16.972 | 18 | 115 | 149 | 1 | 3 |
| 23 | 21.308 | 20 | 126 | 151 | 2 | 4 |
| 114 | 15.867 | 11 | 122 | 152 | 1 | 3 |
| 16 | 15.962 | 13 | 128 | 162 | 1 | 3 |
| 156 | 20.417 | 14 | 124 | 148 | 2 | 3 |
| 39 | 16.681 | 21 | 118 | 157 | 2 | 4 |
| 94 | 10.044 | 10 | 94 | 152 | 2 | 3 |
| 163 | 11.887 | 13 | 99 | 154 | 2 | 3 |
| 99 | 19.664 | 12 | 101 | 151 | 2 | 3 |
| 34 | 13.109 | 11 | 109 | 159 | 1 | 3 |
| 173 | 12.606 | 8 | 105 | 171 | 1 | 3 |
| 66 | 13.448 | 7 | 87 | 157 | 1 | 4 |
| 90 | 18.893 | 21 | 147 | 155 | 2 | 4 |
| 165 | 9.664 | 7 | 84 | 166 | 1 | 3 |
| 64 | 19.299 | 17 | 109 | 158 | 2 | 3 |
| 144 | 20.751 | 16 | 104 | 161 | 2 | 3 |
| 96 | 21.059 | 24 | 127 | 152 | 2 | 3 |
| 102 | 21.987 | 23 | 111 | 159 | 2 | 4 |

Grade Seven (Continued)

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|--------|--------------|-----|-------------|
| 59 | 20.796 | 19 | 121 | 151 | 2 | 4 |
| 175 | 21.505 | 19 | 129 | 152 | 2 | 4 |
| 121 | 11.687 | 9 | 89 | 159 | 2 | 3 |
| 57 | 15.693 | 7 | 98 | 162 | 1 | 2 |
| 8 | 8.489 | 13 | 99 | 160 | 2 | 2 |
| 159 | 25.604 | 24 | 141 | 153 | 2 | 4 |
| 91 | 19.842 | 21 | 119 | 155 | 2 | 4 |
| 14 | 18.451 | 19 | 102 | 158 | 2 | 3 |
| 117 | 19.715 | 18 | 100 | 157 | 2 | 3 |
| 82 | 12.304 | 11 | 116 | 151 | 2 | 3 |
| Means: | 17.725 | 16.07 | 114.07 | 155.09 | | 3.27 |

2. GRADE NINE:

| | | | | | | |
|-----|--------|----|-----|-----|---|---|
| 158 | 24.946 | 25 | 145 | 172 | 2 | 4 |
| 88 | 15.266 | 12 | 99 | 179 | 2 | 3 |
| 166 | 18.468 | 24 | 120 | 177 | 1 | 3 |
| 18 | 14.697 | 24 | 107 | 181 | 1 | 3 |
| 118 | 24.959 | 25 | 112 | 177 | 2 | 3 |
| 83 | 21.608 | 16 | 109 | 182 | 1 | 3 |

Grade Nine (Continued).

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|-------|--------------|-----|-------------|
| 125 | 12.161 | 7 | 96 | 181 | 1 | 1 |
| 17 | 19.590 | 17 | 96 | 170 | 1 | 3 |
| 168 | 12.425 | 12 | 85 | 188 | 1 | 3 |
| 122 | 16.495 | 14 | 95 | 188 | 1 | 4 |
| 30 | 19.859 | 21 | 122 | 175 | 2 | 4 |
| 170 | 25.892 | 21 | 106 | 183 | 1 | 3 |
| 126 | 28.581 | 27 | 143 | 179 | 2 | 4 |
| 21 | 25.226 | 29 | 147 | 171 | 2 | 4 |
| 76 | 20.048 | 17 | 105 | 173 | 1 | 3 |
| 123 | 19.996 | 22 | 115 | 177 | 2 | 3 |
| 167 | 13.665 | 11 | 105 | 181 | 1 | 4 |
| 112 | 19.609 | 13 | 103 | 181 | 1 | 3 |
| 41 | 21.251 | 22 | 115 | 176 | 2 | 3 |
| 116 | 26.214 | 27 | 120 | 175 | 2 | 4 |
| 111 | 20.456 | 21 | 114 | 178 | 2 | 4 |
| 171 | 12.578 | 12 | 86 | 195 | 2 | 2 |
| 113 | 22.344 | 18 | 121 | 176 | 1 | 4 |
| 6 | 18.562 | 19 | 106 | 182 | 2 | 4 |
| 134 | 20.306 | 23 | 122 | 188 | 2 | 3 |
| 131 | 21.815 | 9 | 125 | 177 | 1 | 5 |

Grade Nine (Continued).

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|-------|--------------|-----|-------------|
| 22 | 19.345 | 18 | 119 | 177 | 2 | 3 |
| 127 | 19.147 | 9 | 118 | 178 | 1 | 4 |
| 148 | 21.263 | 23 | 105 | 173 | 1 | 4 |
| 174 | 21.781 | 13 | 109 | 174 | 1 | 5 |
| 51 | 24.050 | 24 | 114 | 171 | 2 | 4 |
| 133 | 7.426 | 6 | 102 | 182 | 1 | 3 |
| 172 | 19.989 | 19 | 121 | 176 | 2 | 4 |
| 15 | 14.200 | 8 | 100 | 179 | 1 | 4 |
| 176 | 14.449 | 10 | 98 | 181 | 1 | 3 |
| 128 | 13.299 | 13 | 91 | 180 | 1 | 4 |
| 155 | 21.319 | 23 | 113 | 181 | 2 | 4 |
| 137 | 21.606 | 23 | 105 | 175 | 2 | 3 |
| 7 | 11.553 | 6 | 91 | 184 | 1 | 3 |
| 97 | 16.166 | 17 | 94 | 185 | 2 | 3 |
| 19 | 20.936 | 20 | 118 | 180 | 2 | 3 |
| 164 | 21.078 | 18 | 122 | 179 | 2 | 4 |
| 49 | 12.608 | 11 | 105 | 175 | 2 | 3 |
| 4 | 20.919 | 12 | 126 | 173 | 1 | 4 |
| 130 | 23.210 | 17 | 123 | 172 | 1 | 5 |
| 52 | 8.623 | 10 | 86 | 181 | 1 | 2 |

Grade Nine (Continued)

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|--------|--------------|-----|-------------|
| 42 | 17.029 | 15 | 97 | 185 | 1 | 3 |
| 135 | 16.271 | 24 | 105 | 177 | 2 | 3 |
| 69 | 14.063 | 8 | 102 | 173 | 1 | 4 |
| Means: | 18.721 | 17.04 | 109.86 | 178.6 | | 3.45 |

3. GRADE ELEVEN:

| | | | | | | |
|-----|--------|----|-----|-----|---|---|
| 92 | 18.382 | 23 | 123 | 194 | 1 | 3 |
| 67 | 25.447 | 28 | 143 | 206 | 1 | 4 |
| 29 | 22.975 | 29 | 122 | 196 | 2 | 4 |
| 100 | 26.774 | 21 | 129 | 190 | 1 | 4 |
| 2 | 21.862 | 24 | 116 | 203 | 2 | 2 |
| 93 | 21.257 | 26 | 145 | 201 | 2 | 4 |
| 31 | 25.257 | 24 | 111 | 205 | 2 | 3 |
| 89 | 14.835 | 19 | 104 | 217 | 1 | 3 |
| 120 | 15.182 | 13 | 112 | 199 | 1 | 3 |
| 24 | 25.334 | 28 | 126 | 194 | 1 | 3 |
| 162 | 22.289 | 22 | 113 | 199 | 1 | 3 |
| 85 | 27.049 | 17 | 101 | 200 | 2 | 3 |
| 60 | 10.280 | 11 | 110 | 207 | 1 | 2 |

Grade Eleven (Continued).

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|-------|--------------|-----|-------------|
| 193 | 26.670 | 29 | 158 | 193 | 1 | 4 |
| 28 | 22.173 | 25 | 108 | 212 | 1 | 3 |
| 53 | 18.735 | 26 | 122 | 200 | 2 | 3 |
| 84 | 16.588 | 14 | 115 | 197 | 1 | 4 |
| 54 | 29.228 | 30 | 122 | 196 | 2 | 5 |
| 157 | 11.672 | 11 | 95 | 205 | 1 | 4 |
| 95 | 22.671 | 18 | 110 | 197 | 1 | 4 |
| 58 | 26.806 | 22 | 119 | 201 | 2 | 4 |
| 139 | 20.049 | 18 | 120 | 203 | 2 | 3 |
| 47 | 22.661 | 23 | 125 | 197 | 1 | 4 |
| 87 | 21.790 | 18 | 120 | 196 | 2 | 2 |
| 199 | 17.712 | 20 | 103 | 199 | 2 | 3 |
| 187 | 22.910 | 25 | 128 | 187 | 2 | 3 |
| 149 | 19.641 | 18 | 119 | 198 | 2 | 3 |
| 65 | 20.250 | 18 | 104 | 204 | 2 | 3 |
| 154 | 21.959 | 12 | 107 | 211 | 1 | 3 |
| 10 | 19.940 | 22 | 124 | 200 | 1 | 3 |
| 46 | 17.961 | 22 | 115 | 201 | 1 | 3 |
| 36 | 25.060 | 24 | 109 | 203 | 1 | 4 |
| 74 | 20.865 | 24 | 105 | 205 | 1 | 4 |

Grade Eleven (Continued)

| Subject Number | "Complete" Score | Writing Score | I. Q. | Age (Months) | Sex | Self Rating |
|----------------|------------------|---------------|--------|--------------|-----|-------------|
| 150 | 27.245 | 21 | 114 | 205 | 1 | 4 |
| 104 | 25.540 | 24 | 157 | 184 | 1 | 3 |
| 50 | 24.061 | 23 | 123 | 195 | 1 | 3 |
| 107 | 24.832 | 25 | 142 | 200 | 2 | 4 |
| 147 | 22.870 | 18 | 116 | 195 | 2 | 3 |
| 56 | 20.416 | 28 | 118 | 203 | 2 | 4 |
| 77 | 25.237 | 27 | 110 | 195 | 1 | 3 |
| Means: | 21.812 | 21.75 | 119.07 | 199.82 | | 3.35 |

APPENDIX B

An analysis of variance was performed using the combined scores for each transformation for each class in order to determine whether overall developmental trends were significant, scores being percentage of acceptable replies for various items on the "complete and delete" test.

Summary of Analysis of Variance

| Source of Variation | SS | DF | MS | F | P |
|-----------------------------------|-------------|----|---------------|-------|----------|
| Between Subjects: | - | 37 | | | |
| 'A' (Transformation) Main Effects | 370,854.187 | 3 | 3,123,618.062 | 1.451 | 0.246 |
| Subjects Within Groups | - | 34 | 85,179.125 | | |
| Within Subjects: | 483,072.0 | 76 | | | |
| 'B' (Grades) Main Effects | 185,351.875 | 2 | 92,675.94 | 40.48 | 0.000005 |
| A-B Interaction | 54,996.250 | 6 | 9,166.04 | | 0.0018 |
| Bx Subject Within Groups | 151,088.000 | 68 | 2,282.21 | | |

APPENDIX C

Spearman Rank Order Correlations For Groups of

Transformations Across Grades

1. Embedding Transformations

| Transformation | Rank/Grades | | | 7x9 | | 7x11 | | 9x11 | |
|----------------|-------------|------|------|-----|----------------|------|----------------|------|----------------|
| | 7 | 9 | 11 | d | d ² | d | d ² | d | d ² |
| 1 | 1 | 1 | 1 | | | | | | |
| 2 | 2 | 2 | 2 | | | | | | |
| 3 | 3 | 3 | 3 | | | | | | |
| 4 | 4 | 5 | 6 | 1 | 1 | 2 | 4 | 1 | 1 |
| 5 | 5.5 | 4 | 4 | 1.5 | 2.25 | 1.5 | 2.25 | | |
| 6 | 5.5 | 6 | 6 | .5 | .25 | .5 | .25 | | |
| 7 | 7.5 | 7 | 8 | .5 | .25 | .5 | .25 | 1 | 1 |
| 8 | 7.5 | 9 | 9 | 1.5 | 2.25 | 1.5 | 2.25 | | |
| 9 | 9 | 10.5 | 6 | 1.5 | 2.25 | 3 | 9 | 3.5 | 12.25 |
| 10 | 10 | 10.5 | 10 | .5 | .25 | | | .5 | .25 |
| 11 | 11 | 8 | 11 | 3 | 9 | | | 3 | 9 |
| 12 | 12 | 12 | 12 | | | | | | |
| 13 | 13 | 13 | 13.5 | | | .5 | .25 | .5 | .25 |
| 14 | 14 | 16 | 15 | 2 | 4 | 1 | 1 | 1 | 1 |
| 15 | 15 | 17 | 18 | 2 | 4 | 3 | 9 | 1 | 1 |
| 16 | 16 | 14.5 | 16 | 1.5 | 2.25 | | | 1.5 | 2.25 |

Embedding Transformations (Continued).

| Transformation | Rank/Grades | | 7x9 | | 7x11 | | 9x11 | | |
|------------------------|-------------|------|------|----------------|------|----------------|-------|----------------|---|
| | 7 | 9 | d | d ² | d | d ² | d | d ² | |
| 23 | 17 | 14.5 | 13.5 | 2.5 | 6.25 | 3.5 | 12.25 | 1 | 1 |
| 24 | 18 | 18 | 17 | | | 1 | 1 | 1 | 1 |
| Sums of d ² | | | | 25.00 | | 41.50 | | 30.00 | |

N = 18

(Note - for the calculation below, figures of "r" significant at the .01 level are indicated *; at the .05 level, **.)

The Formula for Pearson's r is:

$$r = 1 - \frac{6\sum d^2}{N(N^2-1)}$$

For 7x9, $r = 1 - \frac{6(25)}{18(323)} = 1 - \frac{150}{5814} = 1 - .026 = .974 *$

For 7x11, $r = 1 - \frac{6(41.5)}{18(323)} = 1 - \frac{249}{5814} = 1 - .043 = .957 *$

For 9x11, $r = 1 - \frac{6(30)}{18(323)} = 1 - \frac{180}{5814} = 1 - 0.031 = .969 *$

2. Simple Transformations

| Transformation | Rank/Grades | | | 7x9 | | 7x11 | | 9x11 | |
|------------------------|-------------|---|----|-----|----------------|------|----------------|------|----------------|
| | 7 | 9 | 11 | d | d ² | d | d ² | d | d ² |
| 26 | 1 | 1 | 1 | | | | | | |
| 33 | 2 | 2 | 4 | | | 2 | 4 | 2 | 4 |
| 34 | 3 | 5 | 3 | 2 | 4 | | | 2 | 4 |
| 35 | 4 | 4 | 5 | | | 1 | 1 | 1 | 1 |
| 36 | 5 | 3 | 2 | 2 | 4 | 3 | 9 | 1 | 1 |
| 37 | 6 | 6 | 6 | | | | | | |
| 38 | 7 | 7 | 7 | | | | | | |
| Sums of d ² | | | | 8 | | 14 | | 10 | |

N = 7

For 7x9, $r = 1 - \frac{6(8)}{7(48)} = 1 - \frac{48}{336} = 1 - .14 = 0.86 **$

For 7x11, $r = 1 - \frac{6(14)}{7(48)} = 1 - \frac{84}{336} = 1 - .25 = 0.75 **$

For 9x9, $r = 1 - \frac{6(10)}{7(48)} = 1 - \frac{60}{336} = 1 - .17 = 0.83 **$

3. Position Shift Transformation

| Transformation | Rank/Grades | | 7x9 | | 7x11 | | 9x11 | | |
|------------------------|-------------|-----|-----|------|----------------|------|----------------|-----|----------------|
| | 7 | 9 | 11 | d | d ² | d | d ² | d | d ² |
| 27 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| 28 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 29 | 3 | 5 | 5 | 2 | 4 | 2 | 4 | .5 | .25 |
| 30 | 4 | 3.5 | 4 | .5 | .25 | | | .5 | .25 |
| 31 | 5 | 3.5 | 3 | 1.5 | 2.25 | 2 | 4 | .5 | .25 |
| 32 | 6 | 6 | 6 | | | | | | |
| Sums of d ² | | | | 6.50 | | 10.0 | | 2.5 | |

N = 6

For 7x9, $r = 1 - \frac{6(6.5)}{6(35)} = 1 - \frac{39}{210} = 1 - .18 = 0.82$ N.S.

For 7x11, $r = 1 - \frac{6(10)}{6(35)} = 1 - \frac{60}{210} = 1 - .28 = 0.72$ N.S.

For 9x11, $r = 1 - \frac{6(2.5)}{6(35)} = 1 - \frac{15}{210} = 1 - .07 = 0.93$ N.S.

4. Deletion Transformations

| Transformations | Rank/Grade | | | | | | d ² | |
|------------------------|------------|---|----|---|----------------|---|----------------|----|
| | 7 | 9 | 11 | d | d ² | d | | |
| 14 | 1 | 3 | 2 | 2 | 4 | 1 | 1 | |
| 18 | 2 | 2 | 4 | | | 2 | 4 | |
| 19 | 3 | 1 | 1 | 2 | 4 | 2 | 4 | |
| 20 | 4 | 4 | 5 | | | 1 | 1 | |
| 21 | 5 | 6 | 6 | 1 | 1 | 1 | 1 | |
| 22 | 6 | 5 | 3 | 1 | 1 | 3 | 9 | |
| Sums of d ² | | | | | | | 10 | 20 |

N = 6

For 7x9, $r = 1 - \frac{6(10)}{6(35)} = 1 - \frac{60}{210} = 1 - .28 = 0.72$ N.S.

For 7x11, $r = 1 - \frac{6(20)}{6(35)} = 1 - \frac{120}{210} = 1 - .57 = 0.43$ N.S.

For 9x11, $r = 1 - \frac{6(10)}{6(35)} = 1 - \frac{60}{210} = 1 - .28 = 0.72$ N.S.

APPENDIX D

INTER-MARKER RELIABILITY

The method used is a three-way analysis of variance, and generalizability theory suggested by Lee J. Cronbach, Nageswari Rajaratnam, and Goldine C. Gleser, "Theory of Generalizability: A Liberalization of Reliability Theory," British Journal of Statistical Psychology, XVI (21, November, 1963), 137-163.

ANALYSIS OF VARIANCE-3WAY TABLE

| Source | Sum of Squares | D.F. | Mean Square | F Ratios | Probability |
|-----------|----------------|------|-------------|----------|-------------|
| A-Subject | 798.09 | 133 | 6.00 | 6.19 | 0.00 |
| B-Tests | | 1 | | | |
| AB | 219.67 | 133 | 1.65 | 1.70 | 0.00014 |
| C-Markers | 0.032 | 2 | 0.016 | 0.017 | 0.983 |
| AC | 315.63 | 266 | 1.186 | 1.22 | 0.0503 |
| BC | .366 | 2 | .183 | 0.19 | 0.828 |
| Error | 257.96 | 266 | 0.97 | | |

$$6.00 = .97 + 2(.11) + 3(.23) + 6\sigma_S^2 \quad \sigma_S^2 = 0.69$$

$$0.16 = .97 + 2(.11) + 134(-.59) + 268\sigma_m^2 \quad \sigma_m^2 = 0.29$$

$$0.0 = .97 + 134(-.59) + 3(.23) + 402\sigma_T^2 \quad \sigma_T^2 = 0.19$$

$$0.18 = .97 + 134(\sigma_{MT}^2)$$

$$\sigma_{MT}^2 = -0.59$$

$$1.19 = .97 + 2(\sigma_{SM}^2)$$

$$\sigma_{SM}^2 = 0.11$$

$$1.65 = .97 + 3(\sigma_{ST}^2)$$

$$\sigma_{ST}^2 = 0.23$$

$$\sigma_e^2 = 0.97$$

For Tests, $N = 2$

For Markers, $N = 3$

$$\text{Reliability} = \frac{\sigma_e^2}{\sigma_e^2 + \frac{1}{N_m} \sigma_M^2 + \frac{1}{N_T} \sigma_T^2 + \frac{1}{(N_m)(N_T)} \sigma_e^2}$$

Substituting values into the formula given

$$\text{Reliability} = \frac{.97}{.97 + \frac{.29}{3} + \frac{.19}{2} + \frac{.97}{6}} = 0.73$$

APPENDIX E

***PLEASE DO NOT OPEN THIS BOOKLET UNTIL TOLD TO DO SO.

LAST OR FAMILY NAME _____

FIRST NAMES _____

SCHOOL ATTENDED _____

GRADE _____ BOY _____ GIRL _____
(Check One)

AGE IN YEARS AND MONTHS _____ YEARS, _____ MONTHS

HOW WOULD YOU RATE YOUR OWN ABILITY IN ENGLISH? CIRCLE ONE:

POOR BELOW AVERAGE AVERAGE GOOD VERY GOOD

(Do not write below this line.)

INT: _____

COMP. WR. SCORE: _____

CLOZE SCORE _____

COMPLETE AND DELETE TEST

You are asked to help us in the writing of several stories as follows:

INSTRUCTIONS:

This test involves two procedures:

1. Filling in blanks with SINGLE words so that the sentences fit into the passages or stories and make sense.
2. From the passages or parts of sentences that are underlined, crossing out as many words as you can without changing the sense or meaning of that sentence.

EXAMPLES OF COMPLETION ITEMS:

1. A man shot thedog.

(If you wrote any one word in the blank which described the dog, your answer would be acceptable. e.g. big, black, barking, savage, sick, etc.)

2. I saw the manthe moon.

(The obvious word in the above blank would be "in," but other words would also fit: from, with, throwing, etc.)

EXAMPLES OF DELETION (CROSS OUT) ITEMS:

1. John eats grapes because he likes to eat grapes.
(DELETE)

(If you crossed out the second "eat grapes", your answer would be acceptable. Cross out unnecessary words in either of two ways, either

...he likes to eat grapes. /

or ...he likes to ~~eat/grapes~~. /

2. Sam saw an ice cream truck which was in the park.
(DELETE)

(If you crossed out "which was", your answer would be acceptable. Be sure that you cross out everything you want to cross out, and only what you want to cross out.)

NOTE:

- *Observe PUNCTUATION, and DO NOT add any of your own.
- *Read every sentence completely before filling in blanks.
Put in the word that fits BEST and MOST NATURALLY for you.
- *Work quickly. Try all test items. You can come back if
if necessary.

NOTE: Answers are keyed to the list of transformations in
in chapter three, and answers are merely suggested ones.

SELECTION A: CHARLIE'S CHORE

Charlie had a chore to do./ His chore was tiresome./

..... 37 (It) was boring./ And Charlie, a normal

..... 7 (boy) loathed boring jobs./ 17 (If)

asked what he hated most, 6 (Charlie's) vote would be

for monotonous tasks./ This particular task was hauling out

trash./ Charlie often 18 (said), "I hate

hauling trash!"/

Not that the trash was particularly heavy, mind

you./ No, it was simply 9 (that) the alley was a

long distance away, and Charlie hated 4 (making)

more than one trip./ So naturally, Charlie tried to take

everything in one trip whenever his mother called to him,

".....³⁶ (take).....out the garbage!"/

Charlie could tell by the pitch of his⁶ (mother's)
 voice¹⁷ (whether).....he could stall any longer,
²⁵ (and).....eventually he would give in to his fate./

One Saturday morning, as the echoes of his
⁶ (mother's)..... voice died away, Charlie knew ~~that~~¹⁴ the
 moment had come./ (DELETE)

Dragging his way to the kitchen⁵ (with).....
 little enthusiasm, Charlie surveyed the three large bags
² (which)..... were piled on top of the two step-on cans./
³⁵ (What)..... should he do?/

"That doesn't look too bad,"³⁰ (thought)..... Charlie./
 ".....⁴ (Carrying)..... those should be a snap if I can just get
 the kitchen door open."/ Then he had an idea./ ".....¹⁷ (If).....
 I can prop the door open⁵ (with).....a stick, I can kick
 the stick away as I go through," he thought.³⁷ (It).....
 worked./ His idea of propping open the door seemed to give
 Charlie new strength./ He knew it was vital¹² (for).....

him to arrange his load skilfully./ A bag.....^{5 (with)}
 a wet bottom could easily collapse with disastrous results./
 Charlie^{26 (was)} not going to be beaten by a bag of
 garbage!/
 One bag^{1 (which)} was obviously the oldest

looked rather tired and soggy./ Charlie was^{24 (so)}
 careful with it that a spectator would have^{18 (said)},
 "I'll bet it's full of nitroglycerine!"/ He placed^{37 (it)}
 gingerly on the counter./^{31 (Then)} he arranged the
 other two bags on either side./ After^{22 (that)}, he
 carefully took the handle of one garbage pail in each hand./
 The bags^{1 (which)} remained were gathered lovingly in
 his outstretched arms./ His nose was directly over the
^{3 (middle)} bag, and his nostrils were attacked
^{26 (by)} the smell of old potato peels./ He knew
^{9 (that)} he couldn't take that odor for long, so he
 hastened to stagger for the door and fresh air./

Alas!/.^{12 (his)} propping the door had been a

touch of genius, but²² (now) it became his downfall./
 One foot just cleared the stick, but the other trapped it
 between his legs,⁴ (tripping) him and sending him
 flying, garbage bags, cans, and all, onto the porch and
 half-way down the steps./³⁵ (Was) Charlie mad?/ He
 took it very well, like a professional gambler² (who)
 finally loses./³⁰ (Said) Charlie, "At least it
 didn't happen last Saturday when I had two pails and five
 bags!"/

SELECTION B: MY TIME OF YEAR

You may argue that for me¹¹ (to) like autumn
 best is queer, but that's how I feel./ The¹⁵ (fact)
 that winter is coming is undeniable, but Christmas is also
 coming./²⁷ (There) are also winter sports just around
 the corner./

But autumn has many other features ~~which~~ you cannot
deny./ (DELETE)³⁴ (Can) anyone doubt that autumn is
 the most colorful season?/ You may wonder²³ (where)

you wish, but you cannot escape colorful canopies of

20
leaves ~~which~~ are waiting for Winter's blast./ (DELETE)

Sharp spikes of evergreens push their way up out of billowy

seas of golden poplars./¹⁰ (To)..... see hillsides

ablaze with reds and oranges of smaller bushes ¹⁹is an

experience ~~which~~ few can forget./ (DELETE) Yes,

.¹² (autumn's)..... decorating of the countryside is one of
 the year's memorable experiences./

Autumn has other attractions which appeal to the
 sportsmen, however./ My father, a ⁷ (hunter)....., loves
 to walk the fields with his shotgun ²⁵ (and)..... a dog in
 search of pheasants or partridge ² (which)..... rocket into
 the air in heart-stopping blurs./ It is especially exciting
¹⁶ (when)..... several birds take off at once./ Big
 game hunting is another pastime which appeals to sportsmen./

.....⁸ (That)..... deer are a challenge to hunt is an

undisputed fact./ ²⁰A deer ~~which~~/is jumping through bush

(DELETE) is a ³ (challenging)..... target./ There are
 an

^{33 (not)}
 many thrills as heart-stopping as bringing
 down a running deer./

Autumn is also a time of harvest ^{23 (when)} you
 reap the fruits of the year's labor./ Farmers are happy
¹⁴
to know ~~that~~ their efforts (DELETE) have ^{33 (not)}
 been in vain./ There is no doubt that the farmer's hardest
 toil of the year is represented by ^{13 (his)} autumn
 harvesting; but bringing the crops ^{29 (in)} is a labor
 of love, a high ^{7 (point)} in the year./ However,
^{32 (when)} I think of harvesting grain, I personally
 feel itchy all over./ I remember ¹⁹ one autumn ~~when~~ I drove a
combine (DELETE) for three dusty weeks./ After ^{22 (that)}
 I was always busy ^{16 (when)} some farmer friend needed
 help./

^{27 (There)}
 are so many other reasons for
^{13 (my)} loving autumn that I could go on forever./
 However ^{32 (whenever)} I consider carefully, I suppose my
 primary reason is ^{11 (that)} I somehow get a feeling of

settling down, of moving ahead with some purpose to my life
in fall after the lazy summer./

SELECTION C: OLD SHORTY

"Most unforgettable characters" are often quickly
forgotten^{26 (by)} others who only read a description./
But the^{15 (fact)} that Shorty was a character would be
obvious to anyone seeing him./^{17 (Whether)} you believe
it or not, in this day when everyone wears a belt, he wore
fireman's suspenders a full three inches wide./ Shorty
seemed proud of^{38 (himself)} and his suspenders./ He
would walk around giving^{28 (them)} a tug from time to
time, or hooking his thumbs in his trousers top to stretch
them up and down nervously./ He would^{18 (say)}, "Yep,
suspenders don't put a pot on you./^{27 (There)} are too
many folks, gents and gals too, as sticks their stomachs
.....^{29 (out)} to hold up their trousers."/

He would go^{23 (when)} he wanted a drink to a
local hotel, and find a chair against a wall./ Why against

a wall?/ Because when he drank, he didn't like anyone getting behind him./ Even when he was outdoors, he would often say, ".³⁶ (Don't) injun up on me!"/

He was seldom in the city, however, ²¹ because he didn't like it ~~in/the/city~~./ (DELETE) Shorty took good care of ³⁸ (himself), and he hated the smoke and noise of urban life./ "Show me a city man," said ³⁰ (Shorty) often,

"and ²⁰ I'll show you a man ~~who/is~~ breathin' poison!"/ (DELETE)

He often ¹⁴ said ~~that~~ he would sooner be a hermit (DELETE) than have to live in a city forever./

How did Shorty spend his time, you ask?/ He was a prospector, and ²⁴ (so) good at his job that he was in demand./ When working, he always gave his ²⁸ (employer) his best./

³⁴ (Did) Shorty have any advice for young people, you ask?/ He always said, "³⁶ (imagine) you're always workin' for yourself, and then don't disappoint the boss!"/

SELECTION D: PEACE IN OUR TIME

The nations of the world find themselves in a strange situation with respect to the promotion of peace./ The powerful nations do³³ (not) trust anybody./ The nuclear powers¹ (who) hold the fate of the world in their grasp dare not risk total involvement in any armed clash./⁸ (That) they are caught in a web of fear is obvious./³⁴ (Will) the situation change?/ Is there hope for¹³ (our) achieving understanding?/

It is not certain⁸ (that) people want peace./ We solve one armed conflict,²¹ and another ~~armed/conflict~~ (DELETE) takes its place./ Hatred seems a part of human nature./ The¹⁵ (fact) that conflict never seems to cease is undeniable./³⁵ (What) is the answer?/¹⁰ (To) solve the riddle is a³ (great) an necessity because of modern technology./ Never³¹ (before) has man been capable of destroying³⁸ (himself) and the rest of the world as well./ A madman's¹² (pushing).....

a button could be the end./ We have moved^{24 (so)}.....

close to the brink that even militarists admit fear./

.....^{31 (Without)}..... doubt,^{10 (for)}..... us to

survive, we must change./ We must change our sense of

values and ²¹~~we must change our~~ attitudes toward others./

(DELETE) We must give^{28 (them)}..... the benefit of the

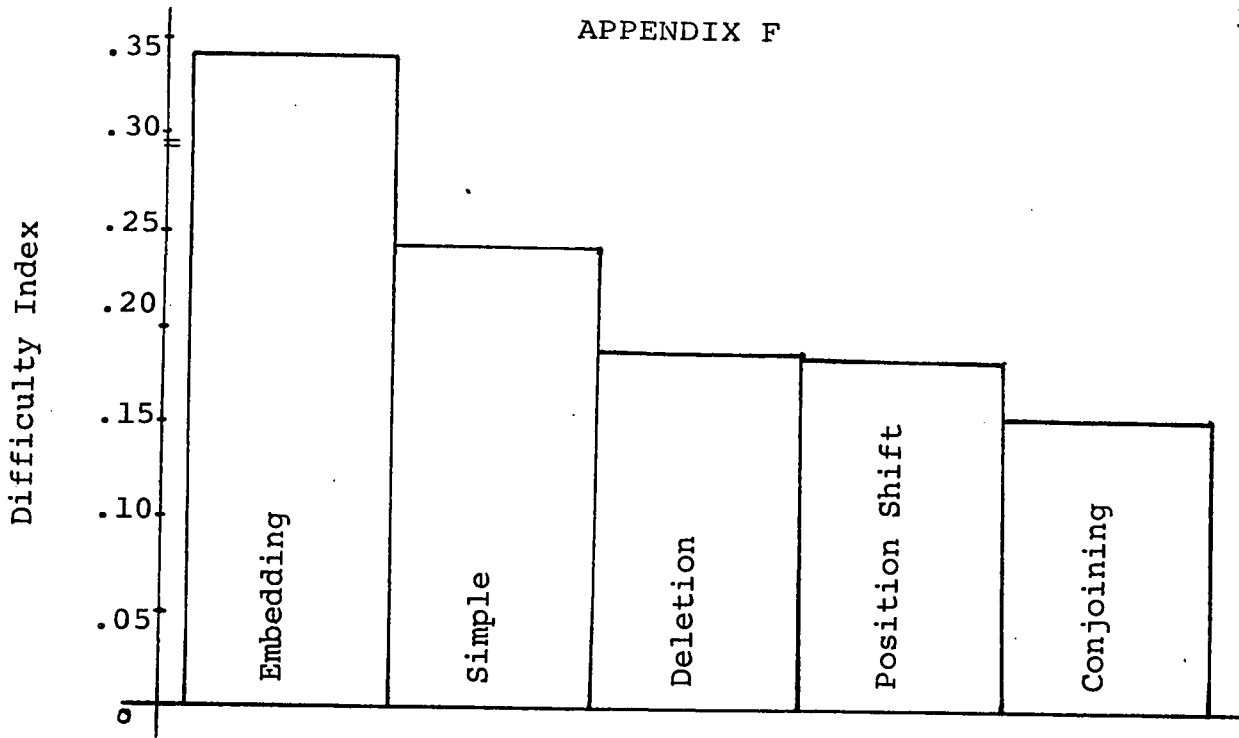
doubt./^{32 (When)}..... we can all learn to trust others,

then we can all live without fear. But a world of hatred

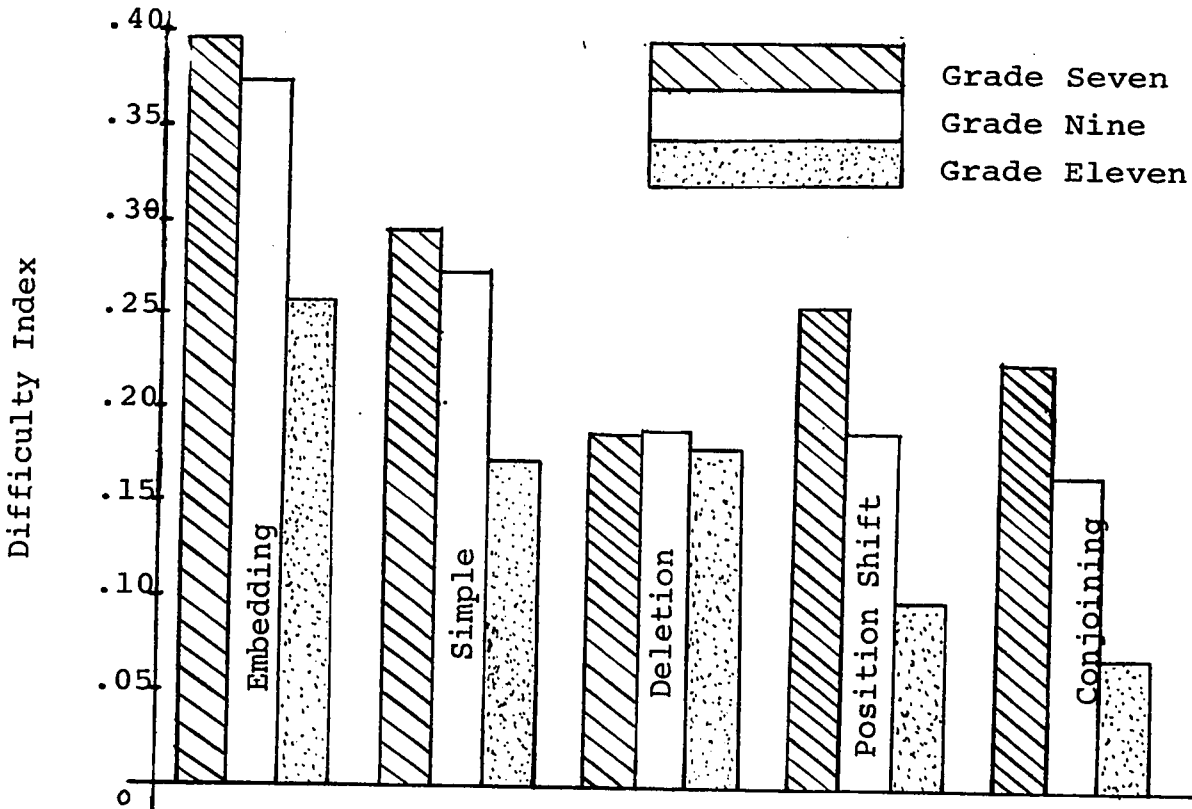
and distrust is a world of growing terror, a time bomb

ticking the seconds^{29 (away)}..... ./ Could it be

.....^{9 (that)}..... our time on Earth will run out?/



Transformation Difficulty Indices Overall



Transformation Difficulty Indices by Grades