INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
Neufeld, Jacqueline Krause

The Effects of New-Word Density on Measures of Reading Comprehension in University English as a Second Language Students

The Ohio State University

University Microfilms International

Copyright 1980 by Neufeld, Jacqueline Krause

All Rights Reserved
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ⌇.

1. Glossy photographs ________
2. Colored illustrations ________
3. Photographs with dark background ________
4. Illustrations are poor copy ________
5. Print shows through as there is text on both sides of page ________
6. Indistinct, broken or small print on several pages ________ throughout ________
7. Tightly bound copy with print lost in spine ________
8. Computer printout pages with indistinct print ________
9. Page(s) ________ lacking when material received, and not available from school or author ________
10. Page(s) ________ seem to be missing in numbering only as text follows ________
11. Poor carbon copy ________
12. Not original copy, several pages with blurred type ________
13. Appendix pages are poor copy ________
14. Original copy with light type ________
15. Curling and wrinkled pages ________
16. Other ________________________________
THE EFFECTS OF NEW-WORD DENSITY ON MEASURES OF READING COMPREHENSION IN UNIVERSITY ENGLISH AS A SECOND LANGUAGE STUDENTS

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Jacqueline Krause Neufeld, B.A., M.A.

******

The Ohio State University

1980

Reading Committee: Approved By

Professor Gilbert A. Jarvis
Professor Victor M. Rental
Professor Betty Sutton

Advisor
To Rocky, for his patience.
ACKNOWLEDGMENTS

I wish to thank Dr. Gilbert Jarvis for all the help, guidance, and encouragement he has given me throughout the course of my doctoral program. I particularly wish to thank him for the endless hours he spent in editing this dissertation. I would also like to thank Dr. Victor Rental for all the time he spent in advising me through the intricacies of reading research; and, to Dr. Betty Sutton for her expert advice and suggestions in the field of English as a Second Language. A special note of thanks to JoAnn Ard, Roger Stuebing, and Don Walizer for their patience, encouragement, and helpful criticisms of this dissertation as it progressed through its various stages.
VITA

December 17, 1948 ............ Born—Columbus, Ohio

1970 ......................... B.A., The Ohio State University
Columbus, Ohio

1972 ......................... M.A., University of Cincinnati
Cincinnati, Ohio

1973 ......................... Graduate studies—Université de Paris (Pantheon), Paris, France

1974-1978

Instructor, English as a Second Language; Instructor, Reading
and Study Program, University of Cincinnati, Cincinnati, Ohio

1978-1979 ..................... Director, English as a Second
Language, University of Cincinnati, Cincinnati, Ohio

PUBLICATIONS

"ESL Teacher Training: A Plea for a Complete Curriculum." in
to Speakers of Other Languages, 1979, pp. 239-249.

FIELDS OF STUDY

Major Field: Foreign Language Education in Teaching English
as a Second Language
FIELDS OF STUDY—Cont.

Studies in Foreign Language Research and Testing.
Professor Gilbert A. Jarvis

Studies in English as a Second Language Methodology.
Professor Edward Allen

Studies in Linguistics. Professor Arnold Zwicky, The Ohio State University; Professors Joseph Foster and William Lasher at the University of Cincinnati
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. iii
VITA .......................................................... iv
LIST OF TABLES ................................................................. ix
LIST OF FIGURES ................................................................. x

Chapter

I. THE PROBLEM ................................................................. 1

    Introduction to the Problem ......................... 1
    Theoretical Bases ........................................ 4
          Cloze as a test of reading comprehension .... 8
          Nonsense words in reading research .......... 10
    Statement of the Problem .................................. 11
    Operational Definitions .................................. 12
    Value of the Experiment .................................. 15
    Limitations of the Study ................................ 16

II. REVIEW OF THE LITERATURE ........................................... 18

    Overview ..................................................... 18
    Nature of Language and Meaning ....................... 19
    Psycholinguistics .......................................... 23
    Vocabulary and Reading Comprehension ............... 28
    Nonsense Syllables and Comprehension ............... 32
    Cloze as a Test of Reading Comprehension ......... 35
    English Language Readability Studies ............... 40
    Summary ..................................................... 45

III. DESIGN AND PROCEDURES ............................................. 46

    Population and Sample ................................. 46
    Design of the Experiment ............................... 47
    Procedures .................................................. 50
    Materials .................................................... 50
    Instrumentation ........................................... 57
    Pilot Study .................................................. 58
### III. DESIGN AND PROCEDURES--Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results of the Pilot Study</td>
<td>59</td>
</tr>
<tr>
<td>Discussion</td>
<td>67</td>
</tr>
</tbody>
</table>

### IV. RESULTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>71</td>
</tr>
<tr>
<td>Hypotheses Tested</td>
<td>74</td>
</tr>
<tr>
<td>Statistics Indicating a Linear Trend</td>
<td>80</td>
</tr>
<tr>
<td>Results Obtained by Means of Scheffé</td>
<td>83</td>
</tr>
<tr>
<td>Summary</td>
<td>85</td>
</tr>
</tbody>
</table>

### V. DISCUSSION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>87</td>
</tr>
<tr>
<td>Research Questions and the Results</td>
<td>87</td>
</tr>
<tr>
<td>Theoretical Bases Reviewed</td>
<td>95</td>
</tr>
<tr>
<td>Discrepancies: Pilot &amp; Experimental Studies</td>
<td>104</td>
</tr>
<tr>
<td>Summary</td>
<td>108</td>
</tr>
</tbody>
</table>

### VI. CONCLUSIONS, SUMMARY & OVERVIEW, AND IMPLICATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusions</td>
<td>110</td>
</tr>
<tr>
<td>Summary and Overview</td>
<td>111</td>
</tr>
<tr>
<td>Design</td>
<td>111</td>
</tr>
<tr>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>Implications</td>
<td></td>
</tr>
<tr>
<td>Recommendations and Limitations</td>
<td>116</td>
</tr>
</tbody>
</table>

### BIBLIOGRAPHY

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIBLIOGRAPHY</td>
<td>118</td>
</tr>
</tbody>
</table>

### APPENDIXES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Instructions for Reading Selection</td>
<td>128</td>
</tr>
<tr>
<td>B. Instructions for Cloze Measure</td>
<td>129</td>
</tr>
<tr>
<td>C. Fry's Readability Graph</td>
<td>130</td>
</tr>
<tr>
<td>D. Description of Subpopulation--Pilot</td>
<td>131</td>
</tr>
<tr>
<td>E. Analysis of Variance and Orthogonal Polynomial Regression Statistics</td>
<td>132</td>
</tr>
</tbody>
</table>
## APPENDIXES--Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>F. Stepwise Regression Summary Table--Pilot</td>
<td>134</td>
</tr>
<tr>
<td>G. Description of Subpopulation</td>
<td>135</td>
</tr>
<tr>
<td>H. Analysis of Variance and Orthogonal Polynomial Regression Statistics</td>
<td>136</td>
</tr>
<tr>
<td>I. Stepwise Regression Summary Table</td>
<td>137</td>
</tr>
<tr>
<td>J. 95% Confidence Levels</td>
<td>138</td>
</tr>
<tr>
<td>K. Scheffe for the .05 Level</td>
<td>139</td>
</tr>
<tr>
<td>L. Scheffe for the .100 Level</td>
<td>140</td>
</tr>
<tr>
<td>M. Reading Passage: Control 1--Scrambled</td>
<td>141</td>
</tr>
<tr>
<td>N. Reading Passage: Treatment 1 (1:2)</td>
<td>143</td>
</tr>
<tr>
<td>O. Reading Passage: Treatment 2 (1:4)</td>
<td>144</td>
</tr>
<tr>
<td>P. Reading Passage: Treatment 3 (1:6)</td>
<td>146</td>
</tr>
<tr>
<td>Q. Reading Passage: Treatment 4 (1:8)</td>
<td>148</td>
</tr>
<tr>
<td>R. Reading Passage: Treatment 5 (1:10)</td>
<td>150</td>
</tr>
<tr>
<td>S. Reading Passage: Treatment 6 (1:12)</td>
<td>152</td>
</tr>
<tr>
<td>T. Reading Passage: Treatment 7 (1:14)</td>
<td>154</td>
</tr>
<tr>
<td>U. Reading Passage: Control 2--Unaltered Text</td>
<td>156</td>
</tr>
<tr>
<td>V. Cloze Test of Comprehension</td>
<td>158</td>
</tr>
<tr>
<td>W. Nodine &amp; Hardt's Histogram of P. Lat.'s for 2100 CVC Trigrams</td>
<td>161</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1. Cloze Test Criterion Scores Equivalent to Traditionally Accepted Performance Levels</td>
<td>38</td>
</tr>
<tr>
<td>2. Illustration of the Design of Experiment</td>
<td>51</td>
</tr>
<tr>
<td>3. Description of Subpopulation--Pilot</td>
<td>60</td>
</tr>
<tr>
<td>4. Analysis of Variance and Orthogonal Polynomial Regression Statistics; Breakdown for Step One--Pilot</td>
<td>62</td>
</tr>
<tr>
<td>5. Stepwise Regression Summary Table--Pilot</td>
<td>64</td>
</tr>
<tr>
<td>6. Independent Variable Coaded to a Percentage and its Equivalent Ratio</td>
<td>73</td>
</tr>
<tr>
<td>7. Description of Subpopulation</td>
<td>76</td>
</tr>
<tr>
<td>8. Analysis of Variance and Orthogonal Polynomial Regression Statistics; Breakdown for Step One</td>
<td>81</td>
</tr>
<tr>
<td>9. Stepwise Regression Summary Table</td>
<td>84</td>
</tr>
<tr>
<td>10. Cloze Test Criterion Scores Equivalent to Traditionally Accepted Performance Levels</td>
<td>88</td>
</tr>
<tr>
<td>11. Expected Score on Cloze Test from Percentage of Known Words in Selection</td>
<td>96</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure                                      Page

1. Linear Characterization of the Means--Pilot............ 66
2. Linear Characterization of the Means .................... 75
3. 95% Confidence Intervals ................................. 78
4. Illustration of the Slope and Intercept of Line .................. 86
5. Diagram of Nonsense Words inserted According to Ratio ......................... 92
Chapter I

The Problem

**Introduction to the Problem**

Educators in the field of second-language acquisition have long agreed upon the importance of developing reading fluency in their students. Reading fluency naturally presupposes a measurable level of comprehension; this, then, makes comprehension of the passage the single most important factor in reading. What a reader tries to do when reading is to reconstruct a coded message; failure to reconstruct this coded message results in a lack of comprehension. Because of the emphasis upon understanding the written message, educators have been wary of introducing too many unfamiliar words into reading selections. Although only one factor among many, it has long been assumed that comprehension is adversely affected by the introduction of too many unfamiliar words. To date, however, there has been very little research in this area. This study, therefore, will examine the effects of new-word density on measures of reading comprehension with
university-level, English as a Second Language students.

The importance of vocabulary to the reading process is not a new theme in reading research. Klare (1963) stated that "vocabulary is the key variable in reading comprehension" (p. 20). Tests and research in native-language studies have shown that vocabulary correlates very highly with reading and mental development (o'Rourke, 1973). Ruddell (1968) concluded that reading comprehension is a result of the reader's control over not only the lexical but also the grammatical components of written discourse.

Many questions related to the role of new-word density on reading comprehension remain unanswered. Clearly, much work needs to be undertaken to elucidate the effect of new-word density on the communication process for pedagogical purposes; some questions that need answering are:

1. What is the ratio of new words to unfamiliar words that interferes with comprehension of the written passage?

2. What is the ratio of new words to unfamiliar words that interferes with the comprehension of oral speech?

3. What is the optimal point at which new vocabulary words could be introduced into the reading material for effective learning to take place?
Before questions of new-word density ratios can be tested, however, the nature of the reading passage must be taken into account. Because comprehension is the single most important factor in reading, a determination of the difficulty level of the passage becomes a necessary pre-step in this type of research. In determining the readability of any passage, the two most important factors that must be considered are the reader and the materials. It was not until recently, however, that techniques devised to tap reading comprehension considered the readers—a fact that doomed earlier techniques to failure.

Trying to eliminate the subjective problems inherent in using intuition to determine readability, reading specialists in the past have devised formulas to judge the difficulty level of the materials for their students. Not entirely satisfied with this approach, some reading specialists began to determine the difficulty of the passage by directly testing the students through the use of comprehension questions. For their respective times, these two techniques were a marked improvement over teacher intuition. They are, however, now thought to be severely limited in several areas. One very important area of concern is their low reliability and validity coefficients as test instruments. Even Fry (1977), the
author of "Fry's Readability Graph," acknowledged that readability formulas are not strong in reporting either validity or reliability (p. 243). But there are other aspects, too, which cause concern over the use of these two techniques. Nelson (1978) cautioned that readability scores gained by the use of these techniques do not deal with "levels of abstraction, complexity of concepts, variations in format and organization, and multiple meanings" (pp. 620-625). Nelson's criticism focused upon the failure of these techniques to account for important linguistic factors.

Bormuth (1968) accounted for the disregard of crucial linguistic factors by citing the then poorly defined theories in the fields of linguistics and psychology. Because of the recent advance in these areas, much more is known about the crucial factors involved in reading comprehension. With this knowledge came insights into how to test for comprehension and for readability. The readability method that most closely utilizes the insights gained from psycholinguistic research is the cloze format.

Theoretical Bases

Psycholinguistics. The theoretical underpinning for this study rests upon the psycholinguistic reading model. While psycholinguistics does not offer a method for teaching reading,
it does offer a theory of how the reader comprehends the written word. Psycholinguistics is a mixture of two previously distinct sciences: cognitive psychology and linguistics. Cognitive psychology views learners as active rather than passive participants in the learning process: this process enables learners to test hypotheses about the learning situation and, based upon their knowledge of the world and feedback, either confirm or reject their hypotheses. In other words, for the cognitive theorist, learners have the capability of expanding their mental structures through the use of logic, reasoning, and increasingly powerful learning strategies (Ausubel, 1978; Smith, 1973). Transformational-generative linguistics classifies language into two components: surface and deep structure. The surface structure of a language can be viewed as the production of auditory phonemes or as the actual ink marks on a written page. Meaning, on the other hand, is generated by means of the deep structure (Chomsky, 1965).

The psycholinguistic reading model can be conceptualized as an information processing system in that it details how visual information is processed. Visual information on a page, according to this model, is picked up by the eye after having been directed to do so by the brain. As the information begins coming in, the brain processes it for meaning.
using its syntactical knowledge of the language to enable it to make predictions, as well as using the content of the information. In other words, the bridge between meaning (deep structure) and visual configurations on a page (surface structure) is the syntax. The frame of reference to make the message relevant and meaningful for the reader is the knowledge of the world which the readers carry about in their heads (Cooper & Petrosky, 1976).

For Smith (1973), reading is communication in which the readers bring with them all their prior knowledge and experiences in order to make sense out of the reading act. Comprehension in reading is a matter of sampling surface structure or the visual information of language in order to test predictions and resolve uncertainty about underlying meaning. Similarly, Goodman's (1968) model of the reading process is based upon three cueing systems: grapho-phonemic (sound-symbol), syntactic, and semantic. Goodman believes that readers reconstruct a message that has been encoded by a writer as a graphic display. If they are fluent readers, they do this in a direct manner; they sample, and rely on the redundancy of the language and on their knowledge of linguistic constraints. They predict structures, test them against the semantic context that they construct from the situation and from interaction with the
discourse, and then confirm or disconfirm as they process further language.

In establishing the theoretical importance of psycholinguistics for second-language learning, Yorio (1971) has reduced the process described above to the following elements:

1. knowledge of the language (code),
2. ability to predict in order to make correct choices,
3. ability to remember the previous cues,
4. ability to make the necessary associations between the different cues selected.

Yorio then modified these factors to reflect more accurately this process with the second-language reader:

1. The reader's knowledge of the foreign language differs from that of the native speaker.
2. The guessing or predicting ability necessary to pick up the correct cues is hindered by his imperfect knowledge of the language.
3. Wrong or uncertain choices of cues make associations difficult.
4. Lack of training and unfamiliarity with the materials make it more difficult to remember cues previously decoded.
5. At all times and levels there is interference (pp. 108-109).
The disadvantages for the second-language reader are several, as Yorio described them: a) second-language readers sometimes either recall cues imperfectly or do not recall them at all, and b) they have a tendency to forget what went before as they are processing subsequent information (p. 109). If psycholinguists are correct in their estimation of the reading process, it can be assumed that with the introduction of too many unfamiliar vocabulary words the process of prediction will fail. With the failure of the prediction system and the inability to make accurate hypotheses, it is further assumed that comprehension will be impeded.

**Cloze as a test of reading comprehension.** The development of psycholinguistics not only has offered a model of how the written word is comprehended, but, due to its insights into the reading process, has also spurred new interest in the cloze test as a measure of determining comprehension. Originally, the cloze test was designed by Taylor (1953) as a measure of readability, but researchers quickly began to use it to test comprehension. A problem the early researchers had to overcome, when using the cloze as a measure of comprehension, was how to interpret the results. In fact, Walter (1974) stated, "The problem of how to interpret cloze test scores has plagued researchers more than any other facet of the cloze procedure" (p. 23).
Because cloze tests may vary in length, use of raw scores was meaningless. Converting raw scores to percentages did not yield any more information on how well the student actually comprehended a passage. Historically, however, reading teachers and researchers, when using a multiple choice test for comprehension, have designated two points on a hundred point scale as measures of student performance. If a student scored 90% on a multiple-choice test, it was assumed that the reading selection could be read by the student without supervision. In other words, the student could read the passage independently of his/her instructor. If the student scored 75% on a reading selection, it was assumed that the material could be read only with supervision—-with the help of an instructor. The problem that faced the researchers in the 1960's was to find a cloze criterion score for each of these two traditionally accepted performance levels.

Through a series of experiments, Bormuth (1967c) matched cloze test results with performance on multiple-choice comprehension tests. Replicating his own experiment the next year, Bormuth (1968) established cloze scores that could be equated to the performance levels traditionally held for multiple-choice comprehension tests. In addition to these levels, he also established a level for beginning comprehension to take
place, which he called frustration level. Frustration level was defined as the beginning of comprehension, a level at which the student feels frustrated and discouraged. Several researchers have since replicated Bormuth's studies with similar results. In defining new-word density ratios and their relationship to and effect upon comprehension, this study will use the performance levels that were established by Bormuth. They are:

<table>
<thead>
<tr>
<th></th>
<th>Multiple-Choice</th>
<th>Cloze</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent</td>
<td>90%</td>
<td>57%</td>
</tr>
<tr>
<td>Instructional</td>
<td>75%</td>
<td>44%</td>
</tr>
<tr>
<td>Frustration</td>
<td>-</td>
<td>35% (p. 62).</td>
</tr>
</tbody>
</table>

Nonsense words in reading research. When designing an experiment in which the effect of increasing new words on reading comprehension is examined, it is crucial to eliminate any variables that might confound the results. Psychologists and psycholinguists have shown how people learn: through using prior information in order to make the learning task meaningful. In an experiment, however, if a person uses prior information to answer a question correctly, the results of the experiment will be confounded. Ever since Ebbinghaus' (1885) famous experiments using nonsense syllables, psychologists
have used the nonsense syllable when they wanted a word that would be completely independent of meaning. Researchers have tested nonsense syllables and some have been found to be low in orthographic, phonologic, and semantic meaningfulness. In using these nonsense syllables, psychologists, therefore, are assured that the person being tested is not using prior knowledge to answer the questions.

When testing English as a Second Language subjects, one can never know for certain the exact composition of each person's lexicon. Because these English as a Second Language students come from a variety of backgrounds, it would be unwise to assume that they shared a common vocabulary. In such a case the only practical solution in testing the effect of new-word density on reading comprehension is to isolate a group of students who is matched within a range in their language ability, and then to use a reading selection that has been chosen specifically for that range of ability.

Statement of the Problem

It is the purpose of this study to examine the effects of increasing the number of nonsense syllables on measures of reading comprehension for intermediate-level students of English as a Second Language at an American university. The following research questions will be considered:
1. What is the effect of increasing the ratio of nonsense syllables on measures of reading comprehension for intermediate-level English as a Second Language students?

2. What is the ratio of nonsense syllables to English words that results in frustration level comprehension? (35% correct on cloze measures; Bormuth, 1967c)

3. What is the ratio of nonsense syllables to English words that results in instructional level comprehension? (44% correct on cloze measures; Bormuth, 1967c)

4. What is the ratio of nonsense syllables to English words that results in independent level comprehension? (57% correct on cloze measures; Bormuth, 1967c)

**Operational Definitions**

The following operational definitions will be used in preparing materials for this study.

1. **Comprehension**

   The term *comprehension* is considered in the study to be a construct measured by means of a cloze passage administered to the subject upon completion of the reading of the stimulus passage. The following criteria serve to define operationally the performance levels of comprehension measured by the cloze passage.
a) **frustration level**: the point at which comprehension of a passage begins. Although beginning comprehension is not a discrete variable, Bormuth (1967c), Oller (1971), and Rankin (1975) view 35% correct on cloze measures as the point at which comprehension of a passage begins.

b) **instructional level**: the point at which students can read material without frustration, with the help of an instructor. Bormuth (1967c), Oller (1971), and Rankin (1975) view 44% correct on cloze measures as the minimum necessary for the instructional level.

c) **independent level**: the point at which students can read materials independent of instructional support. Bormuth (1967c), Oller (1971), and Rankin (1975) view 57% correct on cloze measures as the minimum necessary for the independent level.

2. **Nonsense Syllable**

The term *nonsense syllable* is used here to mean any word which is low in orthographic, phonologic, and semantic meaningfulness and which does not, therefore, resemble English.
3. **Stimulus Passage**

The _stimulus passage_ is defined here to mean a passage which, when subjected to the cloze test for readability, will result in intermediate English as a Second Language students scoring at least 44%.

4. **Intermediate English as a Second Language Students**

_Intermediate English as a Second Language Students_ will be defined in this study by use of the mean, the standard deviation, and the placement procedures of a sample population for both the Michigan Test of English Language Proficiency and the Test of English as a Foreign Language (TOEFL). The average of several sample means and standard deviations on the Michigan Test was 75% and 12% respectively (Interpretive Manual, 1965). The average of several sample means and standard deviations on the Test of English as a Foreign Language was 500 and 70 respectively (Interpretive Manual, 1978). The students in this study will be defined, therefore, as having scores between 76% and 89% on the Michigan Test of English Language Proficiency or 500 and 570 on the Test of English as a Foreign Language. Placement suggestions for academic work for both tests suggest that students who score within
these ranges may take up to 75% of a normal load with one to two English courses, depending on their score within the range.

**Value of the Experiment**

Traditionally, foreign language educators have had to rely on intuition and conjecture when determining the readability of reading material. It is believed that: a) this study will reveal the point at which the introduction of unfamiliar vocabulary words affects the comprehension of the reading passage; b) the results of this study will provide empirical data that will have implications not only for reading comprehension but for vocabulary development as well; c) the formula which will be obtained through the statistical analysis will permit English as a Second Language instructors to predict the performance levels for their intermediate-level students when applied to reading material like that used in this experiment.

Psycholinguists assert that if information is related during the reading process to existing knowledge in the person's cognitive structure, meaningful learning will occur. Psycholinguistics has analyzed what fluent readers do as they read, but, as yet, there has been little empirical research to validate these models. Until these models can be expanded
upon through knowledge gained from research, pedagogical practices will continue to rely on intuition.

The importance of reading in second-language learning can not be overstressed. Perhaps among all the linguistic skills that a second-language learner will gain in the course of studying the target language, reading is the skill with which s/he will develop the greatest ability. It is the skill, moreover, that will endure when the others have been forgotten. For these reasons, research into reading comprehension constitutes a significant endeavor (Chastain, p. 308).

Limitations of the Study

Because of the nature of meaningful learning and of reading comprehension, an undertaking of both these topics in any one study would end in futility. It is the purpose of this study to show, specifically, the point at which the reader will begin to understand (comprehend) an unfamiliar reading passage. Translated into classroom practices, this research question will attempt to answer the question of how many new/unfamiliar words are too many for comprehension to take place. Of course, the whole question of the effect of new-word density on comprehension is irrelevant if meaningful vocabulary learning does not take place. This study, however, will not attempt to answer the question of how many new words can be
meaningfully learned in context; rather, it will restrict itself to determining what ratio of unfamiliar to familiar words will impede comprehension. As a by-product of the experiment, this study will attempt to establish new-word ratios for independent and instructional reading levels.

This does not mean, however, that readability of the passage will be introduced as a variable. The material, chosen for intermediate students, represents a difficulty level directed for instructional reading by intermediate students. The question of readability and new-word density ratios could be an entire study in itself. It is only after this necessary first step that further investigations can be undertaken to answer questions dealing with readability and meaningful vocabulary learning.
Chapter II

Review of the Literature

Overview

This study will attempt to determine the effect of new-word density upon comprehension for English as a Second Language readers. Because most comprehension studies with English as a Second Language subjects have based their underlying assumptions upon native-language comprehension studies, it is first necessary to discuss the theoretical bases that underlie language studies in general. By discussing these language universals, certain assumptions about the second-language learner can then be reasonably drawn. The author, as well as some psychologists, psycholinguists, and linguists, believes that all natural languages share certain assumptions, as will be explained in the first section, "Nature of Language and Meaning." The theoretical assumptions (Part I) are used as a rationale for establishing the applicability of psycholinguistics in explaining reading behavior, regardless of language differences, which are discussed separately in Part II. Part III
surveys research in the area of vocabulary and comprehension, and discusses how these two factors are interrelated. This section emphasizes psychological and psycholinguistic research and points a direction for further study. Part IV discusses the use of nonsense syllables in psychological and psycholinguistic research, specifically detailing the theory for their use. Parts V and VI discuss the cloze test as a valid, reliable measure of readability and comprehension. An historical overview is given to show how the cloze format was developed in response to the information gained in early psycholinguistic research.

**Nature of Language and Meaning**

It is because natural languages tend to share some basic characteristics called universals, that one is able to examine psycholinguistic research done with English-speaking subjects and generalize to those whose native language is not English. An example of this would be the knowledge that surface structure differs from language to language, but all natural languages have some kind of order imposed on their surface structure; hence knowledge that is transferrable. Another such language universal is the notion of competence and performance.

Language competence was first defined by de Saussaure (1959) and later refined by Chomsky (1965). Competence is the largely unconscious intrinsic knowledge which underlies a man's ability
to speak and to understand what is spoken. Performance, on the other hand, is the utilization of this intrinsic knowledge, applied in specific cases. One of the simplest explanations of this dichotomy is found by looking at child language acquisition. Within the relatively short time of three to five years, a child will construct a complete grammar of the languages s/he hears based upon exposure to a finite number of utterances in that language (Chomsky, 1965). This innate ability is what Chomsky termed competence and is characteristic of all natural language learners.

The competence/performance model of language acquisition assumes a number of things: a) that the child will actively develop language, and b) will do so based upon an innate language mapping system. The acquisition of language by children develops due to their interactions with their environment through hypothesis testing and feedback. Research with English as both a native language and a second language seems to support this hypothesis (Ervin-Tripp, 1964; Dulay and Burt, 1972; Brown, 1973).

Natural languages share not only the idea of competence/performance, but also another important characteristic dealing with the nature of language itself. This characteristic is that, universally, more abstract structures than surface
structures are defined in languages. Linguistically, these "deep" structures are established for a number of reasons: a) deep structure facilitates syntactic description by making it possible to capture generalizations about the language that are not apparent in the surface structure b) underlying structures reflect semantic properties of sentences that are obscured in the surface structure. This knowledge about language includes the following:

1. synonymy resulting from identity deletion
2. synonymy resulting from the deletion of indefinite elements
3. ambiguity
4. generative ability
5. recognition of ungrammaticality (Chomsky, 1965; Langacker, 1972)

Some linguists, however, believe that language is a mere outward manifestation of an underlying biological characteristic that all persons possess. Lenneberg (1967), in Toward a Biological Theory of Language Development, stated that "cognitive function is a more basic and primary process than language" (p. 374). As with linguists, some cognitive psychologists, notably Ausubel (1978), adhere to Lenneberg's theory that man derives meaning from written and spoken discourse ...
because of his penchant for organization. Highly developed animals such as man will organize their sensory world by a process of categorization. It is from this organization that two further processes derive: a) differentiation or discrimination, which is vital to the fluent reading process, and b) interrelating categories (transformations), by which we form concepts (p. 378). From this theory, meaning is seen to derive through our system of organization rather than through an arbitrary convention exhibited by some people but not by others.

According to Lenneberg (1967):

The biological properties of the human form of cognition set strict limits to the range of possibilities for variations in natural languages. The forms and modes of categorization, the capacity for extracting similarities from physical stimulus configurations or from classes of deeper structural schemata...are powerful factors that determine a particular type of form for language. Within the limits are, however, there are infinitely many variations possible. Thus, the outer form of languages may vary with relative freedom, whereas the underlying type remains constant (p. 375).

Ultimately, what concerns both theories is the way in which ambiguity is resolved in an attempt to derive meaning. Since classroom learning is largely concerned with the acquisition and eventual application of meaning, it is necessary to describe and understand precisely how meaning is derived by the student. Psycholinguistics offers such an explanation.
Psycholinguistics

An integral part of not only reading but also psycholinguistics is concerned with the nature of comprehension. Comprehension has been defined differently by a number of leading theorists. Fries (1963) identified three layers of language meanings that must be accounted for in reading comprehension. These meanings include: a) meaning carried by the lexical items b) meanings carried by the grammatical structure, and c) socio-cultural meanings (p. 7). In trying to define what comprehension is, many reading theorists talk in terms of components as did Pavlak (1974). For him, reading comprehension is a complex process and is affected by the following: a) thinking b) attitude c) interest d) sentence structure e) listening ability f) rate g) vocabulary h) questioning and purpose setting (p. 3). But perhaps the most elegant and the most closely in tune with a modern psycholinguistic theory of reading comprehension is Thorndike's (1917) definition of reading comprehension as "reasoning." Spache (1966), commenting on this lack of consensus by the reading theorists, wrote that "the meaning of comprehension differs from user to user but the term is used glibly" (p. 54).

For reading theorists, reading is not a simple mechanical skill nor a completely visual process (Miller, 1968; Palmer,
1974). Reading is essentially a thought process that one uses to derive meaning from written discourse. Smith (1976) defined comprehension as the absence of uncertainty. For Smith, reading is a questioning-from-feedback process which enables the reader to make predictions; comprehension is answering those questions. From the findings of several well-known researchers such as Goodman (1968), Miller (1968), Venezky (1974), and Smith (1976) emerge three themes, basic to the way in which reading theorists view the reading process:

1. Only a small part of the information necessary for reading comprehension comes from the written page.
2. Comprehension can precede the identification of individual words.
3. Fluent reading is not decoding to spoken language.

Reading theorists are concerned with delineating the process that enables the reader to derive meaning from the written page. The factors or linguistic intuitions involved in this process include the notion of grammaticality. This knowledge enables the reader to see a string of words and determine whether it is well-formed or ungrammatical as in the following:

1. The bull looks terrifying.
2. The bull looks flower.
Our knowledge of the grammar of a language enables us to predict what should follow the verb of a sentence; it is this knowledge that excludes the possibility of the deviant sentence *2. Another intuition that all speakers of the same language share is the understanding of grammatical relations. Intuitively we know that these two sentences,

3. John is eager to please.
4. John is easy to please.

although very similar in their surface structure are not structurally synonymous. A syntactic description of a language must also account for a third intuition, the ability to recognize synonymy regardless of transformations. Thus, a speaker or reader knows that

5. The dog bit the little girl.
6. The little girl was bitten by the dog.

mean essentially the same thing. Finally, a speaker must be able to distinguish ambiguity as in:

7. Visiting relatives can be a nuisance (Chomsky, 1965; Slobin, 1971).

Linguistic competence explains why a reader is implicitly able to understand and identify these factors. Linguists, psychologists, and reading theorists such as Goodman, Miller, Slobin, and Smith, to name a few, feel that each reader brings
to the reading task his/her knowledge of not only these linguis-
tic factors, but also his/her knowledge of the world. The reader uses this knowledge to make predictions, to resolve ambiguity, and, finally, to derive meaning. His/her ability to make predictions, therefore, is based more on the non-visual information than on the actual visual marks on the page. As stated earlier (Chapter I), it is the syntax as well as our knowledge of the world that enables us to look at the visual information on a page and understand its meaning. In other words, comprehension is a matter of sampling the surface structure in order to test hypotheses, make predictions, and resolve uncertainty.

Research into the psycholinguistic model of reading has shown that the reader, as a hypothesis tester, uses various strategies to resolve uncertainty. Olshavsky (1977) identified ten reading strategies common to fluent readers. This particular study not only lent credence to the psycholinguistic theory that readers do utilize strategies, but also identified types of strategies as typical for problem solving. In another study with English as a Second-Language subjects, Brownscombe (1977) concluded that second-language readers do have strategies for reading and understanding structures. An important outcome from this study demonstrated that second-language
readers have a tendency to simplify. Not only is there a language universal proposition which states that languages change through simplification; but Goodman (1968) and Smith (1976) predicted that native-language readers sample only enough information to confirm their hypotheses; again, a language simplification process. A famous study by Shannon (1950) seems to confirm the prediction theory. He stated, "The new method of estimating entropy exploits the fact that anyone speaking a language possesses, implicitly, an enormous knowledge of the statistics of the language" (p. 54). According to Shannon, our use of redundancy reduces the number of possibilities of what will come next in a reading passage. His empirical research supports the psycholinguistic concept that a reader uses the language's redundancy to reduce uncertainty.

The underlying assumption is that natural languages of the world share common features, universals. It would seem a reasonable assumption that any universal in human behavior must, ultimately, reflect something fundamental about the organization of the human mind (Goodman, 1968). Much research and speculation on second-language reading mirrors this view. Stephens (1977) hypothesized that a student who reads his/her native language does not need to learn a theoretical system of grammar rules in the traditional sense because s/he possesses
the ability to recognize words in terms of the linguistic factors already mentioned. Stephens further hypothesized that same ability is the foundation for reading in the target language. Chomsky (1964), in a summary of generative-transformational grammar, supports the idea of the applicability of underlying linguistic factors to all natural languages.

The grammar as a whole can thus be regarded, ultimately, as a device for pairing phonetically represented signals with semantic interpretations, this pairing being mediated through a system of abstract structures generated by the syntactic component. Thus, the syntactic component must provide for each sentence (actually, for each interpretation of each sentence) a semantically interpretable deep structure, and a phonetically interpretable surface structure, and, in the event that these are distinct, a statement of the relation between these two structures...Roughly speaking, it seems that this much structure is common to all theories of generative grammar, or at least compatible with them (p. 52).

Once this process of relating surface structure to deep structure is learned in the native language, psycholinguists have speculated that the ability to transfer the knowledge to the second language would develop without having to understand the fundamentals of the process with each new language. The process will remain the same while the structures, vocabulary, and familiarity will change.

Vocabulary and Reading Comprehension

Most researchers and teachers agree on the importance of vocabulary in reading comprehension. Klare (1963) stated that
vocabulary is the key variable in reading comprehension; it is also the most important part of most tests related to academic aptitude and achievement (p. 16). Nearly ten years later Twaddle (1973) asserted that although vocabulary is a very important component of language mastery, there is mostly "agreement in disagreement" concerning the best technique for teaching its usage (p. 62). Spache (1966) said, "Understanding the vocabulary is second only to the factor of reasoning in comprehension..." (p. 78). These and many other researchers have concentrated on validating the concept that vocabulary and reading are highly correlated. What has not been dealt with extensively is the effect of increasing the ratio of new words upon measures of reading comprehension (Holley, 1973; Marks, 1974).

An interesting study, which could prove helpful in understanding how people learn vocabulary words was conducted by Henning (1973). His study dealt with two proficiency levels (high and low) of second-language learners. His results indicate that second-language students encode vocabulary in acoustic and semantic memory clusters. He found that low proficiency learners tended to register meanings in memory more by sound similarities than by relating meanings to something already known. High proficiency learners relied on associated
meanings rather than on sound similarities. This idea of increasing vocabulary through meaningful learning has been dealt with by others as well.

In recent research, Holley (1973) found that by restricting structure and content with her students she was able to introduce as many as one new word for every 15 known words into a reading passage. This, she claimed, was about twice what had been thought possible. A similar study by Marks, Doctorow, and Wittrock (1974) found that sizable increases in reading comprehension could be achieved with elementary school children if the semantic variable of word frequency was controlled.

MacGinitie (1961) found through experimentation with "cloze-like" deletions in paragraphs that uninterrupted context beyond five words did not help the subjects restore the missing words (p. 127). Since MacGinitie's research is based upon different periodic omission patterns, his results could reflect differences in both the total available context or in the length of uninterrupted context, though he believes that proximity is the key issue. The results of his study indicate a basic limitation in the length of units that can be restored from context. Although readers may not see every word clearly, they generally can not restore groups of words lost from the story, no matter how few and widely separated. Thus, MacGinitie
feels that the redundancy of English acts for restorative purposes primarily with small portions of the written message. If these omitted portions are small enough, they may be restored quite accurately, even though they occur frequently—one cloze space for every four words. These research findings indicate that factors such as context and cues play a more important role in comprehension than does the rate of vocabulary introduction. If this is true, perhaps educators should not try to manipulate the structure of the material so much as try to teach successful hypothesis testing and prediction strategies.

Perhaps now that some preliminary studies have been done, the question of how vocabulary and comprehension interact can successfully be attacked. Eisenhardt (1974) succinctly stated what has not occurred in the field of vocabulary and reading:

What has been neglected in the area of vocabulary is the concept that words alone do not carry precise meaning until they are placed in the structural and international system of the English language. An individual's sense of structure gives the written work meaning or determines how it means in contrast to what it means (p. 1).

A key question that should be determined is how best to apply the psycholinguistic theories in order to teach reading strategies and vocabulary effectively so as to boost comprehension.
Nonsense Syllables and Comprehension

Reading research has revealed that readers use prior information during the reading process; because of this, researchers have relied on nonsense syllables to help achieve as true an estimate of reading comprehension as possible.

Ebbinghaus (1875) first tested nonsense words to judge their effect on recognition, recall, and memory. In similar experiments dealing with perception, Cattell (1885) found that a familiar unit can be reconstructed from a bare minimum of perceptual clues because of the limited number of ways in which ambiguous portions can be filled in. More than half a century later, Shannon (1950) empirically confirmed this hypothesis. When nonsense syllables are inserted in place of meaningful words, every part of the pattern must be correctly perceived. There is little chance that an omitted or ambiguously perceived portion can be supplied by the context alone. This is essentially a statistical statement, because it involves the relative chance of supplying missing parts to both familiar and unfamiliar patterns (Miller, Bruner, and Postman 1954). Shannon (1950), and Miller, Bruner, & Postman (1954) asserted that the context operates to reduce the effective range of alternatives, and thus to reduce the amount of information the symbol carries. For example, if a pattern for
letter occurrence is briefly exposed to the subjects but they see an I and nothing else, their chance of correctly guessing the next letter is one in 26 if they know the pattern is nonsense. If, however, they know that the letter is part of a meaningful word, their best guess would be that the next letter would be N. In English the letter N following I occurs one time in every five. When nonsense syllables are used in the reading selection, readers should be affected in the following ways: a) their prediction strategies will be based only upon the number of meaningful words in the passage; b) because they cannot use prior information, they will not try to relate the nonsense syllables in a meaningful way.

Theios and Muise (1975) identified properties necessary to make nonsense syllables useful in research. According to them, nonsense syllables have to be low in orthographic, phonologic, and semantic meaningfulness. First Noble (1961) and later Nodine and Hardt (1969) experimented with nonsense syllables to determine which best represented these three qualities. Nodine and Hardt (1969) scaled 2100 CVC trigrams for pronunciability by measuring pronunciation latency. Their findings can be interpreted in the following manner: the higher the P Lat. of the trigram, the more free of bias from stimulus attributes such as meaningfulness (p. 211). The P Lat.
can be viewed as reflecting a basic processing time for stimulus materials; the more difficult trigrams took longer to process and reflected a lower ratio of meaningfulness. For example, average P Lat. was .81 seconds which included such trigrams as "REV," "FEM," and "SEX." The following P Latencies reflect longer processing time and, hence, lower meaningfulness: a) 1.009 - "QOX" and "QUJ" b) 1.168 - "VEJ" and "GUQ" c) 1.487 "XOJ" etc.

This trigram research was later used and supported by both Smith and Goodman. According to Smith (1975), one factor which makes the learning task more meaningful is the familiarity of the reader with the subject or words in the passage. Thus, Smith regards nonsense syllables as meaningful to the extent that they resemble sequences of sounds or letters in English. According to this definition, "STR" would be more meaningful than "TSR" because the former can occur in a meaningful sequence in English: the more meaningful, the more easily retrieved (pp. 154-158). Goodman (1975), likewise, said that nonsense that looks like English is easier to handle than jumbles of letters that have no correspondence. Goodman believes that the closer language is to its use, the closer it is to its basic relationship to meaning and, consequently, the easier to handle and learn (p. 625). Based upon what is known
about how people learn and the types of things they do when they process language, it can be assumed that if meaningful words were replaced by nonsense syllables, these syllables could not be used to process and predict.

**Cloze as a Test of Reading Comprehension**

In the recent past, traditional methods and theories of reading have come under attack; part of this attack has been focused upon measures that evaluate reading comprehension. This attack has come from not only reading theorists but also from linguists. Those attacking traditional methods of determining reading comprehension, e.g. multiple-choice questions that deal with literal recall, main ideas, conclusions, inference, etc., feel that asking such questions is too indirect a method for yielding trustworthy results. According to Bormuth (1970), there are two inherent weaknesses in using a multiple-choice test of comprehension: a) these tests are not reliable measures of content difficulty because the questions do not always reflect the difficulty of the language used in the passage, and b) the manner in which the test questions are generated is neither replicable nor is it drawn from a total population of possible questions but reflects the individual biases of the test constructor.
Because of the flaws in item selection and reliability, new interest has been given to the cloze procedure developed by Taylor in 1953. Taylor (1953) was influenced by Osgood, whose theories of learning include the ideas of language redundancy and probabilities. Cloze was originally developed as a measure of readability, but researchers soon found extended use for the system. Rankin (1957) and Bormuth (1965) early discovered the extended use of the cloze test as a measure of reading comprehension. Bormuth (1968) stated: "The cloze procedure offers a valid, convenient, and completely objective method of constructing tests that can be used for measuring reading comprehension abilities or the comprehension difficulties of the passage" (p. 353). By administering two sets of tests, one using the cloze and the other the traditional multiple-choice format, Bormuth (1968) found that while all the traditional concepts of reading comprehension are tested when the cloze format is used, the cloze format offers, in addition, a more reliable test instrument (p. 365). The chief importance, therefore, in using the cloze test over the multiple-choice format is its reliability and validity as a test instrument.

Because the cloze test is so unlike the traditional multiple-choice test of comprehension, reading theorists
initially had problems in interpreting the results. Bormuth (1968) overcame these problems by establishing performance levels on the cloze test of comprehension which could be equated to traditionally held performance levels for the multiple-choice test (p. 365). In this study, Bormuth found that 75% or instructional level was equivalent to 44% on a cloze format, and 90% or independent level was equivalent to 57% on a cloze format (see Table 1).

Bormuth helped establish the importance of the cloze test as a measure of reading comprehension for native English speakers. It was not until 1970 that important results were obtained using the cloze test with English as a Second Language students. Darnell (1970) correlated cloze test scores with scores on the Test of English as a Foreign Language (TOEFL) and reported that the cloze scores correlated .83 with the total scores on the TOEFL. What this showed was that a single cloze test tended to rank students very much the same as a five-part, multiple-choice battery. This multiple-choice battery included measures of structure, vocabulary, listening, reading, and writing. Others namely Spolsky (1968), Stubbs (1974), Aitken (1975), and Oller (1975) have conducted research that supports both Bormuth's and Darnell's findings of cloze as an extremely reliable test of reading comprehension. Oller (1972) stated,
Table 1

Cloze Test Criterion Scores Equivalent to Traditionally Accepted Performance Levels

<table>
<thead>
<tr>
<th></th>
<th>Traditional Scores</th>
<th>Bormuth</th>
<th>Rankin</th>
<th>Anderson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional</td>
<td>75%</td>
<td>45%</td>
<td>44%</td>
<td>41%</td>
</tr>
<tr>
<td>Independent</td>
<td>90%</td>
<td>52%</td>
<td>57%</td>
<td>61%</td>
</tr>
<tr>
<td>Frustration</td>
<td>--</td>
<td>35%</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
"Item analysis and Kr-20 measures of reliability ranging from .80 to .98 indicate considerable internal validity for the cloze test used with non-natives" (p. 2).

Oller (1972) enumerated some of the general characteristics of cloze tests of second-language proficiency. These include:

1. Internal consistency--all of the tests conducted yielded impressive item analysis results; out of 150 items, only seven failed to discriminate significantly (p \( \leq .05 \)).

2. Reliability--in studies that utilize every nth deletion ratio, reliability between .80 and .98 was achieved; this was determined by the kr-20 and half-scores.

3. Validity--cloze scores have yielded correlations with the total score on TOEFL and UCLA ESLPE between .83 and .89.

4. Ease of construction and administration.

5. Standardization of reference criteria (pp. 111-112).

The research to establish the reliability and validity of the cloze test has led to some interesting findings. Flahive (1977) found that the traditional multiple-choice format, when used with his non-native subjects, tested intelligence. His conclusion was that multiple-choice formats test reasoning
ability—the ability to retain facts, to draw conclusions, and to use contextual clues to determine meaning. On the basis of these findings, he questioned the validity of the multiple-choice format. In another separate test, Wilson (1977) found that the cloze test was not biased nor did it differentiate on the basis of language background.

There seems to be a raging controversy, however, over the best method for scoring the cloze passages. Oller (1971), Anderson (1972), and McKenna (1976) all report that synonymous answers with cloze passages in English as a Second Language materials are as good as, if not better than, exact-word scoring. All three report high correlations and little difference in rank as a result of accepting synonymous answers. Rankin (1957), Bormuth (1968), and Ruddell (1968), on the other hand, report that exact-word replacement is necessary to obtain high correlations with the other measures. Bormuth (1968) stated:

This practice of scoring correct just those responses which exactly match the deleted word...is based on findings that including synonyms as correct responses slightly increases the correlations between cloze scores and scores on comprehension tests. But the data show that it does so simply by increasing the variability of the scores, an effect that is far more easily obtained by adding a few items to the test (p. 43).

English Language Readability Studies

A very important factor that is often overlooked in choosing reading passages for students is the difficulty level of
the material. It is important to carefully choose passages and texts so as not to frustrate or lose second-language students. Although the readability factor has long been recognized as important, it has been only in the last 10-15 years that a reliable and valid test instrument to determine passage difficulty was devised. Part of the failure of the earlier techniques in determining passage readability was the exclusion of factors directly related to the learner, such as intelligence, motivation, experience, and aptitude. Instead, the earlier techniques concentrated on the factors inherent in the reading material, such as sentence length and complexity, vocabulary, amount of redundancy, the number of sentences per story, style, purpose, etc. Betts (1949) spoke directly to this two-fold factor of readability when he said:

Reduced to its lowest terms, readability is a two way proposition. The first consideration is the reader—his experience, his interests, his feelings, his motivation, his language facility, his needs, and his reading habits. Any means of predicting readability is valid to the degree that the reader is taken into account. The second consideration is the interest level, the language, the mental constructs, and the mechanical features of the reading material (p. 438).

It was precisely the inability of the early readability formulas to take the reader into account that ultimately decided their failure.
The early researchers concentrated on devising formulas that could be applied easily to any textbook. Unfortunately, the formulas frequently yield different results and fail to measure many factors considered important for readability. Fry (1977), for example, enumerated the necessary elements in determining the difficulty level of a reading passage using his formula:

1. Randomly select three sample passages.
2. Count out exactly 100 words in each; start with the beginning of a sentence.
3. Count the number of sentences in the hundred words, estimating length to nearest 1/10.
4. Average the three passages for sentence length and syllable length (p. 249--see Appendix C).

When Fry's formula is used, it is readily apparent that sentence complexity and conceptual difficulty, not to mention subject factors, are not accounted for. The problems inherent in Fry's formula are by no means singular.

Bormuth (1967a) discussed the weaknesses of readability formulas when he wrote:

Perhaps one of the chief reasons why instructional materials are not routinely evaluated to determine whether they have a suitable level of difficulty is that there has been no technique that is at once convenient, economical, and valid. Readability formulas are convenient, inexpensive, and require
only unskilled clerical assistance to use, but the formulas presently available have validities that range from .5 to only about .7. Moreover, the equations take into account only a limited range of linguistic variables and the variables that are taken into account are, by today's standards, crude (pp. 2-3).

Another method used to determine the readability of a passage was to test the students directly on the information. The students' scores were then measured to some pre-established criterion, most notably 90% and 75%, in order to interpret student performance. Although the direct testing method takes into account both subject and material factors, it is, nevertheless, unreliable for two reasons: a) it is almost impossible to test whether the actual difficulty level of the passage was reflected by the students' scores, and not the difficulty of the test question—which will vary from test writer to test writer; b) a test question should represent an unbiased sample of how well each student comprehended, but a test constructor will actually select the reading passage to be tested based upon personal biases. It is the selection process which reduces the reliability.

Bormuth (1968) attributed the failure of these two techniques to (then) poorly defined theories in psychology and linguistics. Because the theories in these areas were not clearly defined, readability formulas and the direct method
of testing could not illuminate the sources of difficulty in the passage. A method that takes into account the myriad of linguistic factors as well as the subject factors is the cloze test. Taylor (1953) developed the cloze test as a measure of readability based upon the Gestalt psychologist's concept of closure. This concept states that the speaker and listener will anticipate what is coming next in the communicative act and try to "cloze" anything left unfinished. Oller (1972) stated:

The cloze procedure is one of those tests which, like dictation, measures integrative skills, and has much to recommend it. The most important argument in its favor is that it requires the student to perform a task which is not unlike what native speakers do in sending and receiving messages. In listening, we anticipate what we will say next and frequently...supply missing words or phrases. In speaking, we sometimes find ourselves groping for a word half-way into the sentence (p. 187).

In addition to measuring integrative linguistic skills, it is also a measure of the students' ability to comprehend the reading passage. According to Anderson (1976):

The significance of cloze procedure as a measure of readability is that the method appears to take into account both material and subject factors of readability. There is a sound theoretical basis also for cloze procedure as a measure of reading comprehension for cloze scores index the correspondence of language habits of writer to those of reader and thus the construct, reading comprehension, is given operational meaning (p. 7).
Because the cloze tests are constructed by deleting every nth word and are completely replicable, there are none of the inherent weaknesses or biases so frequent in the direct testing method. In addition, because the cloze method reflects the actual difficulty level of the passage, it provides far greater validity and reliability than do either of the other two techniques.

**Summary**

It is clear from the literature relating to comprehension and new-word densities that little has been done in this area. It is also clear from vocabulary studies, in particular, that research has either dealt with correlating vocabulary to comprehension or, separately, determining which vocabulary method worked best for a given sample. The studies show that vocabulary knowledge is highly correlated with reading comprehension; therefore, it should be both logical and valuable to try to discover how these two variables interact during the reading process. Only when this interaction is more clearly understood, will it be possible to address pedagogical considerations. It is the purpose of this study to examine the nature of this interaction.
Chapter III
Design and Procedures

Population and Sample

The sample of this study was drawn from a population of non-native intermediate-level students of English at the University of Cincinnati. About 5% of this population choose studies in the humanities and the social sciences, but the rest were graduate students who choose advanced technological areas of concentration such as the physical sciences, engineering, and mathematics. These students were either foreign aliens or permanent residents who had been in the United States for at least one year but not exceeding three consecutive years. These students came from all five continents, but the majority were from Taiwan and India.

The sample consisted of intermediate English as a Second Language students enrolled at the University of Cincinnati for the academic year 1978-1979. Intermediate level has been operationally defined on the basis of the students' scores on the Michigan Test of English Language Proficiency (76%-89%) and the Test of English
as a Foreign Language (500-570). The ranges from these two tests represent the mean and one standard deviation above the mean; the standard deviation, therefore, is 13 for the Michigan and 70 for the TOEFL. These ranges have been recommended by the test constructors as guidelines for university administrators in placing foreign students with regard to academic load and are an indication that a person can work independently in an American university with minimal support. Minimal support is defined as one to two English classes per quarter in addition to 3/4 of a normal academic load.

**Design of the Experiment**

To maximize information yield and design efficiency, trend analysis using orthogonal polynomials was used to analyze the results in this study. Each student was randomly assigned to one of the treatments illustrated in Table 2. Each treatment consisted of a reading passage and a cloze test for comprehension based upon the passage. The treatments differed in the ratio of nonsense syllables to English words found in each passage; the cloze test, regardless of the treatment, remained the same. The differing ratios in each treatment were arbitrarily chosen to cover the range of possibilities within the stated interval (0:0-1:15). This interval was chosen because
of results from prior research closely related but not identical to this study. In research with American students learning German, Holley (1970) found that an incidence as high as one new word for every 15 known words did not significantly diminish reading comprehension (p. 195). MacGinitie (1961), when studying the effects of contextual constraints upon English prose for native English speakers, found that constraints between words generally decreased very rapidly with distance, and the influence of context upon word choice had relatively little effect when there was a distance of five words or more between cloze blanks. In addition he found that cloze items in which the distance is decreased to four words or less are very difficult to construct (pp. 128-129). In other words, MacGinitie was unable to detect any dependence among items when four or more words of the text remained between cloze items.

Subjects were randomly assigned to one of the various treatments; subject A, for example, was given the reading passage in which the ratio of nonsense syllables to English words was one to two. This meant that for every two English words, one nonsense syllable was inserted to replace a regular English word in the passage. There were seven treatment groups with the following ratios:
In addition there were two control groups which consisted of the following: a) a completely scrambled passage, and b) a rendition of the straight text. These ratios and controls were chosen for the following reasons:

1. No comprehension of the passage is expected when a student reads the completely scrambled passage (control 1), while full comprehension is expected of the student assigned to read the straight text (control 2).

2. If the experiment begins from a point where no comprehension of the passage is expected, and introduces differing ratios of nonsense syllables at decreasing frequencies, then it is assumed that at some point between the two controls comprehension will occur.

3. The ratio of the point at which comprehension occurs will delimit the new-word density which indicates that reading comprehension has been impeded.

The independent variable was the differing ratios of nonsense syllables, while the dependent variable was the score on the cloze test for comprehension. The nonsense
syllables used in this study were randomly selected from the list prepared by Nodine & Hardt (1969; see Appendix W). These researchers discovered which nonsense syllables were low in orthographic, phonologic, and semantic meaningfulness and listed them accordingly; it was from this pool that the nonsense syllables were randomly selected.

**Procedures**

On the day of the experiment each student was randomly assigned to one of the treatments or controls. The students were given a brief introduction to the experiment (see Appendix A). They then read the instructions governing the rules for reading the selection (see Appendix A). Once they had finished reading the selection, they were asked to inform the proctor so that the reading passage could be collected and the cloze test with the appropriate code (corresponding to their treatment) could be assigned. The students were then given another set of instructions (see Appendix B) for reconstructing the cloze items.

**Materials**

Several factors had to be considered in choosing materials for a reading comprehension experiment. These
Table 2
Illustration of the Design of the Experiment

<table>
<thead>
<tr>
<th>Group</th>
<th>Type of Treatment</th>
<th>Scrambled</th>
<th>1:2</th>
<th>1:4</th>
<th>1:6</th>
<th>1:8</th>
<th>1:10</th>
<th>1:12</th>
<th>1:14</th>
<th>Straight Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
included: the difficulty level of the passage, its syntactical characteristics, and its length. The criterion reading selection, "Lincoln's Autobiography", was originally published in American Readings, an intermediate-level English as a Second Language text.

The first consideration was the level of difficulty of the passage. There are two ways in which the difficulty level of a passage can be determined: objectively or subjectively. Objectively, one might take a sample from the same population that is to be tested in the experiment and administer the criterion passage to them. The range of test scores would then indicate the difficulty level of the passage for these students. This method, however, is not feasible at the University of Cincinnati: although the intermediate-level is by far the largest language group of ESL students at this university, they are nevertheless, limited in number. Taking a sample from the same population as the sample for the experiment would have greatly reduced the total number possible for the experiment, and hence the experiment's generalizability. It was thought to be more appropriate, therefore, to judge subjectively the difficulty level of the criterion measure.
The difficulty level was subjectively determined by choosing another selection in the same intermediate-level test, establishing its difficulty level, and then comparing it to the criterion measure. The selection chosen for this purpose was "Among the Osage Indians" also from American Readings. This selection was comparable in length, difficulty, vocabulary, and sentence structure to "Lincoln's Autobiography." The level of difficulty for "Among the Osage Indians" was then determined in the following ways: a) through the use of the cloze test as suggested by Strong (1978) and, b) by means of Fry's (1975) readability formula. An intermediate-level English as a Second Language class was given a cloze test based upon "Among the Osage Indians"; the results of this cloze test showed that the scores for these intermediate-level students fell within the range of 44%-57%, which, desirably, placed the material at the instructional level according to Bormuth (1968). In addition, Fry's readability formula was applied to both selections—"Lincoln's Autobiography" and "Among the Osage Indians." Both selections placed mid-way through the 9th level (see Appendix C). The problem of finding an appropriate difficulty level is crucial to fruitful
reading research dealing with comprehension. If the passage is too difficult or too easy, the scores on the criterion measure will be either all high or all low; no new information could then be deduced from such an experiment. If the results are to provide necessary new information, the passage must represent a difficulty level that intermediate-level English as a Second Language students would normally encounter: that of instructional-level prose.

Another characteristic that had to be considered in looking at the difficulty level of a reading passage was the syntactical complexity of the sentences. Intermediate-level, as opposed to advanced or elementary-level, English as a Second Language texts have certain very distinct syntactical characteristics. Of the four sentence types—simple, compound, and compound-complex—the writer of the intermediate-level English as a Second Language text will choose to intersperse complex, compound, and compound-complex sentences with simple sentences, so as to make the reading more lucid and the focus more direct and easy to find. For instance, in "Lincoln's Autobiography" there were 12 complex sentences, making the complex sentence the predominant type; there were
also four compound and four compound-complex sentences. Interspersed among these complex sentence types were 10 simple sentences. In addition, there were several parenthetical expressions that added to the difficulty of the passage.

In contrast to this, an elementary English as a Second Language reader will have neither complex or compound-complex constructions, nor parenthetical expressions. The reading selections will consist primarily of simple sentences with, perhaps, one or two short compound constructions. At the other end of the scale, an advanced reader will not present materials re-written like the other two types of texts, but will present it in its original form. The reading selections found in advanced readers will consist primarily of compound, complex, and compound-compound constructions. No attempt will be made to intersperse the more complex constructions with simple sentences, as advanced understanding of English by the student is assumed. All the selections in American Readings represent the syntactical characteristics described for intermediate-level English as a Second Language reading selections.

The last significant characteristic in choosing
appropriate reading materials was the length of the passage. In order to obtain a stable score of the difficulty level of the passage, it was necessary to choose a passage which was long enough to delete every fifth word and which would still yield at least 50 blanks (Bormuth, 1968). Taylor (1957) originated the idea of allowing 50 blanks on each cloze test because he felt that this number would be sufficient to yield a stable score of the difficulty level of the passage; sufficient because the chances of mechanically selecting all easy or all hard words, and of selecting words of the same grammatical class, would cancel each other as the length of the text increased (p. 48). Bormuth (1968) later found that the greater the number of blanks on a cloze test, the higher the reliability (p. 50).

"Lincoln's Autobiography" is a short selection of approximately 500 words, which will allow 99 blanks if every fifth word is deleted.

Finally, in order not to confound the results of the measures of reading comprehension, it was felt that titles which could serve as advance organizers should be omitted (Dooling and Lachman, 1971).
Instrumentation

The criterion measure, in the form of a cloze format, was administered to each student after the passage had been read. It was designed to test the individual's reading comprehension of the passage in each treatment. Reading comprehension has been operationally defined by the criterion measure below:

**Cloze passage:** After reading the passage, the student was given a cloze test of reading comprehension which was based on the same passage and in which every fifth word had been deleted (Bormuth, 1968).

The practice of deleting every fifth word is followed because it provides the greatest number of items possible for a given passage and thereby provides the most reliable measure of passage difficulty (p. 48).

Researchers replicating Bormuth's studies with English as a Second Language students (Oller, 1975 et. al) have also determined that every fifth word deletion provided the most reliable estimate of passage difficulty. So as to make the experiment completely replicable, only exact word replacement were tallied. Due to the length of the passage, there were 99 deleted words. Each correctly "clozed" blank, therefore, counted one point.
The cloze passage was designed following Bormuth's (1968) instructions:

1. Randomly select passage(s), if applicable;
2. Always begin at the beginning of a paragraph;
3. Make sample the length of a continuous text;
4. Delete every fifth word (p. 48).

Pilot Study

On April 18, 19, and 20th, 1979, a pilot study was conducted with five sections of the Developmental Reading course at the University of Cincinnati. Students from the Developmental Reading Program were chosen for the following reasons:

1. Developmental American students, none of whom speak standard English, are enrolled in this course in order to develop their English linguistic skills.

2. Studies have indicated that commonalities exist between developmental American students and students learning English as a Second Language (Labov, 1972).

3. Because of the similarities of these two groups, the pilot was run using the developmental students so that no pre-test interaction would result.
with the English as a Second Language subjects.

This study was conducted for the purposes of:

1. refining the criterion measure
2. establishing time considerations
3. becoming better acquainted with the statistical procedures for interpreting the data (trend analysis using orthogonal polynomials)
4. refining the procedures for administering the test

The computer program REGRESSION in the Statistical Package for the Social Sciences (SPSS) was used to analyze the results from the pilot study. The regression analysis, conducted using orthogonal polynomials, investigated the possible linear, quadratic, and cubic effects of "CLARITY" on comprehension as measured by "PCT." The program determined (based upon the results of the experiment) that a linear equation was the best possible characterization of the means (p < .01).

Results of the Pilot Study

The independent variable in this experiment was coded "CLARITY" while the dependent variable was coded "PCT." "CLARITY" measured the total amount of regular English words left in the reading passage after the nonsense
Table 3

Description of Subpopulations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>For entire population</td>
<td></td>
<td>38.18</td>
<td>15.28</td>
<td>233.51</td>
<td>67</td>
</tr>
<tr>
<td>CIARITY</td>
<td>0</td>
<td>23.71</td>
<td>15.75</td>
<td>247.90</td>
<td>7</td>
</tr>
<tr>
<td>CIARITY</td>
<td>67</td>
<td>31.57</td>
<td>15.82</td>
<td>250.29</td>
<td>7</td>
</tr>
<tr>
<td>CIARITY</td>
<td>80</td>
<td>35.25</td>
<td>15.56</td>
<td>242.21</td>
<td>8</td>
</tr>
<tr>
<td>CIARITY</td>
<td>86</td>
<td>43.25</td>
<td>14.25</td>
<td>203.07</td>
<td>8</td>
</tr>
<tr>
<td>CIARITY</td>
<td>89</td>
<td>50.57</td>
<td>17.02</td>
<td>289.62</td>
<td>7</td>
</tr>
<tr>
<td>CIARITY</td>
<td>91</td>
<td>32.75</td>
<td>13.78</td>
<td>189.93</td>
<td>8</td>
</tr>
<tr>
<td>CIARITY</td>
<td>93</td>
<td>41.29</td>
<td>10.70</td>
<td>114.57</td>
<td>7</td>
</tr>
<tr>
<td>CIARITY</td>
<td>94</td>
<td>40.50</td>
<td>14.15</td>
<td>200.29</td>
<td>8</td>
</tr>
<tr>
<td>CIARITY</td>
<td>100</td>
<td>44.86</td>
<td>7.86</td>
<td>61.81</td>
<td>7</td>
</tr>
</tbody>
</table>
syllables had been inserted; "PCT" measured the percentage scored by each student on the cloze test of comprehension administered after the reading. For example, "CIARITY" equalled 67 when 67% of the English words remained in the text; "PCT" equalled 35 when a student scored 35% on the criterion measure.

In looking at the descriptive statistics found on Table 3, one finds a description of the subpopulations, broken down according to the independent variable "CIARITY." The sub-population which has a "CIARITY" of 0 (N=7) has a mean of 23.71 and a standard deviation of 15.74; this was the bottom control group. In the first treatment group, "CIARITY" = 67 (N=7), the mean was 31.57, and the standard deviation was 15.82. In the second treatment group, "CIARITY" equalled 75 (N=8); the mean was 35.25 and the standard deviation was 15.56 (see Table 3 for a breakdown of the descriptive statistics for all seven treatments and two controls; also Appendix D).

A step-wise regression was performed by using the "PCT" scored on the criterion measure. The results of step one are presented below in Table 4. The F ratio is 10.6128 (p < .01) which indicates that the linear effect of clarity accounts for a significant amount of
Table 4

Analysis of Variance and Orthogonal Polynomial Regression Statistics
Breakdown for Step 1

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>PCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables entered on Step 1</td>
<td>CIARITY</td>
</tr>
<tr>
<td>Multiple R</td>
<td>0.37464</td>
</tr>
<tr>
<td>R Square</td>
<td>0.14036</td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.12713</td>
</tr>
<tr>
<td>Standard Error</td>
<td>14.27676</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
<th>d.f.</th>
<th>Sum Squares</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>2163.1730</td>
<td>2163.1730</td>
<td>10.6128</td>
</tr>
<tr>
<td>Residual</td>
<td>65</td>
<td>13248.6777</td>
<td>203.8258</td>
<td></td>
</tr>
</tbody>
</table>

Variables in the equation

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIARITY</td>
<td>0.20168</td>
</tr>
<tr>
<td>Constant</td>
<td>22.3724</td>
</tr>
</tbody>
</table>
variance. In other words, clarity, according to these results, has a linear effect upon comprehension. The multiple R for the linear equation is, in fact, a simple r because there is only a single variable at this step. (see Table 4). The multiple R = .37464 and represents the correlation coefficient of observed to predicted scores. The R square for the linear analysis was .14036 which showed the amount of "PCT" which could be explained by the independent variable, "CLARITY." The standard error, 14.2768, showed the variability of the scores around the predicted linear regression line, expressed in standard units.

Column "B" in Table 4 contains the values of the slope and intercept of the linear equation for these data: PCT = .20 * CLARITY + 22.4. This equation determines the linear regression equation. For example, if a subject receives the control text in which all the English words have been scrambled (CLARITY=0) then his/her predicted score would be 22%; likewise, if a subject receives the straight text in which no permutations occur (CLARITY=100), his/her predicted score would be 42.4% (see Appendix E).
### Table 5

**Step-wise Regression Summary Table**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Multiple R</th>
<th>R²</th>
<th>R² Change</th>
<th>Simple R</th>
<th>B</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity (Linear)</td>
<td>0.37464</td>
<td>0.14036</td>
<td>0.14036</td>
<td>0.37464</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.38520</td>
<td>0.14838</td>
<td>0.00802</td>
<td>0.38520</td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>Cubic</td>
<td>0.38903</td>
<td>0.15135</td>
<td>0.00297</td>
<td>0.36450</td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.7067</td>
<td></td>
</tr>
</tbody>
</table>
Table 5 presents a summary for the entire step-wise regression analysis. Although the F ratio was significant ($p < .01$) for step two, there was relatively little increase in the multiple $R$ from step one to step two (linear to quadratic) and, therefore, the quadratic component was eliminated. Had the multiple $R$ taken a sizeable jump from step one to step two, then it could be defended that the best fit for the data contained a quadratic component in addition to the linear. Because the multiple $R$ was so small from step one to step two, the test for significance on step two (quadratic) ceased to be significant. For graphic illustration, see Table 5. $R^2$ for the linear equation was .14036, for the quadratic .14838, and for the cubic .15135. The amount of change in the $R^2$ from one step to the next can be seen by looking at the next column ($R^2$ change). This represents the difference in change between the respective $R^2$'s. The increase from a linear to a quadratic fit was small: 0.00802, and from a quadratic to a cubic fit was even smaller: 0.00297. This indicates that there is no significant quadratic and cubic relationship between "PCT" and "CLARITY" (see Appendix F).
Figure 1

Linear Characterization of the Means
Discussion

The purpose of the pilot study was to refine both the criterion measure and the administration of the experiment as well as to become better acquainted with the statistical procedures used in the analysis; to these ends, the pilot study was successful. Based upon the pilot study, the following adjustments were made for the present study:

1. Directions had to be made clearer, more precise, and shorter.

2. Even though the directions specifically stated that the subjects were not to look back at the reading passage when filling in the cloze blanks, nevertheless, many covertly tried to find the answers while doing the comprehension test. Because of this, the main study utilized the following format:

   a. Directions for reading the selections (treatments/controls) were included with each appropriate passage. For example, these directions asked the subjects to indicate when they had finished with the reading by holding up their hand. The proctor would
then collect the reading selections from each subject (see Appendix B).

b. The proctor then marked the code for each corresponding treatment or control upon the criterion measure; the criterion measure was then given to each subject with a set of new instructions (see Appendix C).

3. Typographical errors on the cloze test of comprehension were detected and corrected in order to make the study valid. In addition, a decision was reached as to the best possible choice of deleting the blanks on the criterion measure. This was done in order to keep the subjects from correctly guessing 35% or more of the cloze items. In other words, if the subject comprehended nothing but guessed all the function correctly, s/he could still not receive more than 31%; this, then, disallowed anyone from reaching the beginning comprehension mark through guessing alone.

According to the results of this study, the means for the first two groups--control and treatment--remained below the frustration level established by Bormuth (1968);
frustration level being defined as a score between 35%–44% on a cloze test for comprehension. The mean gradually increased by the second treatment (CLARITY=80) in which every fifth word had been replaced by a nonsense syllable. The mean for this group was 35.25%, which marked the beginning of comprehension, but at the frustration level. MacGinitie (1963) had predicted just such a response based upon the results of his research (see Chapter II). The means gradually increased in a linear fashion, then fell with the fifth treatment group before rising again (see Figure 1). The mean for subjects in the second control group (CLARITY = 100) was 44.85%, which should have been the highest of all the groups; however, due to some subjects' not following directions and looking back at the reading selection while answering the cloze questions, there were some very high scores in the middle two sections (CLARITY = 86; CLARITY = 89). Had there been closer controls imposed over these subjects, there is an indication that the means of all the groups would have increased by gradual, but perceptible, degrees. This slight fluctuation in scores did not alter the fact that the linear regression equation was statistically significant.
Through submission of the data to trend analysis using orthogonal polynomials, results indicate the rate at which comprehension changes as the rate of difficulty changes. The following null hypotheses were tested:

1. $H_0$: No comprehension will occur with any of the treatment groups (ratios of nonsense syllables to English words).

2. $H_0$: No comprehension will occur in the control group which receives the scrambled text.

3. $H_0$: No comprehension will occur in the control group which receives the passage with the straight text.
Chapter IV
Results

Introduction

This study examined the effects of increasing the number of nonsense syllables on measures of reading comprehension. Its purpose was to determine if a ratio of nonsense syllables to English words could be established in accordance with predetermined performance levels of reading comprehension (frustration, instructional, and independent). In other words, nonsense syllables were used to determine the ratio of new-word densities that could be tolerated by students while reading; the means from each of the different subgroups were then compared to previously established performance levels so as to judge how well an English as a Second Language student could read, given the varying proportions of a new-word densities for each of the different groups. The results from each group were then submitted to regression analysis using orthogonal polynomials. This analysis investigated the possible linear, quadratic, and cubic effects of the independent variable upon the dependent variable.
The independent variable in this study was the ratio of nonsense syllables to English words and was coded "CIARITY." For the purpose of the analysis, "CIARITY" was transformed into a percent which represented the total percentage of English words remaining in the mutilated passage after the addition of the nonsense syllables. There were nine groups, each with its varying ratios of nonsense syllables to English words; these groups with their equivalent percents and ratios are represented in Table 6. The dependent variable, hereafter referred to as "PCT", was the percent scored by each student on the criterion measure—a cloze test of comprehension.

Subjects in each group were given a reading passage, the only difference among them being the ratio of nonsense syllables inserted in the text. After the subjects had finished the reading, they were given a cloze test of comprehension over the mutilated text. Only correctly "clozed" blanks were counted; scores were then tallied and converted into a percent. The results were submitted to analysis twice: once including those scores from the group which had read the scrambled passage (CIARITY=0), and once without. For reasons which will be thoroughly related in the discussion section, the results with the
<table>
<thead>
<tr>
<th>Percentage of English Text</th>
<th>Ratio of Nonsense to English</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIARITY 0</td>
<td>scrambled</td>
</tr>
<tr>
<td>CIARITY 67</td>
<td>1:2</td>
</tr>
<tr>
<td>CIARITY 80</td>
<td>1:4</td>
</tr>
<tr>
<td>CIARITY 86</td>
<td>1:6</td>
</tr>
<tr>
<td>CIARITY 89</td>
<td>1:8</td>
</tr>
<tr>
<td>CIARITY 91</td>
<td>1:10</td>
</tr>
<tr>
<td>CIARITY 93</td>
<td>1:12</td>
</tr>
<tr>
<td>CIARITY 94</td>
<td>1:14</td>
</tr>
<tr>
<td>CIARITY 100</td>
<td>straight copy of text</td>
</tr>
</tbody>
</table>
data from the control group (CLARITY-0) were discarded. The results section, therefore, is a report on the analysis from the remaining eight groups. The program determined, based upon the results from these eight groups, that a linear equation was the best possible characterization of the means (p<.01—see Figure 2).

Hypotheses Tested

1. \( H_0 \): No comprehension will occur with any of the treatment groups.

Based upon Bormuth's (1968) study on reading performance in which he compared cloze scores to previously established, traditionally held performance levels of reading comprehension (see Table 1, page 32), it is apparent from the means listed in Table 7 that comprehension did occur, in varying degrees, in each of the treatment groups. The null hypothesis, therefore, is rejected.

A "CLARITY" of 67 had a 95% confidence interval of 28.89%-45.11% (\( \bar{X} = 37\% \)); this confidence interval is based upon a passage in which 67% of the English words remained in the mutilated text. According to Bormuth's performance levels for cloze tests of reading comprehension, 35%-44% correct is considered frustration level for reading
Figure 2

Linear Characterization of the Means
Table 7

Description of Subpopulations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>95% Conf. Int.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Variance</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>for entire population</td>
<td></td>
<td></td>
<td>53.46</td>
<td>12.65</td>
<td>160.20</td>
<td>84</td>
</tr>
<tr>
<td>CIARITY 0</td>
<td>0</td>
<td>49.24 - 57.00</td>
<td>53.12</td>
<td>4.64</td>
<td>21.55</td>
<td>8</td>
</tr>
<tr>
<td>CIARITY 67</td>
<td>67</td>
<td>28.89 - 45.11</td>
<td>37.00</td>
<td>11.34</td>
<td>128.66</td>
<td>10</td>
</tr>
<tr>
<td>CIARITY 80</td>
<td>80</td>
<td>46.96 - 60.74</td>
<td>53.85</td>
<td>7.44</td>
<td>55.48</td>
<td>7</td>
</tr>
<tr>
<td>CIARITY 86</td>
<td>86</td>
<td>43.58 - 58.42</td>
<td>51.00</td>
<td>10.38</td>
<td>107.78</td>
<td>10</td>
</tr>
<tr>
<td>CIARITY 89</td>
<td>89</td>
<td>52.01 - 58.43</td>
<td>55.22</td>
<td>4.17</td>
<td>17.44</td>
<td>9</td>
</tr>
<tr>
<td>CIARITY 91</td>
<td>91</td>
<td>50.61 - 66.05</td>
<td>58.33</td>
<td>10.13</td>
<td>102.75</td>
<td>9</td>
</tr>
<tr>
<td>CIARITY 93</td>
<td>93</td>
<td>47.17 - 67.33</td>
<td>57.25</td>
<td>16.51</td>
<td>272.75</td>
<td>12</td>
</tr>
<tr>
<td>CIARITY 94</td>
<td>94</td>
<td>46.47 - 65.13</td>
<td>55.80</td>
<td>13.04</td>
<td>170.18</td>
<td>10</td>
</tr>
<tr>
<td>CIARITY 100</td>
<td>100</td>
<td>49.56 - 70.88</td>
<td>60.22</td>
<td>13.87</td>
<td>192.44</td>
<td>9</td>
</tr>
</tbody>
</table>
comprehension. In other words, students who score within this range are able to comprehend the reading selection only to a small degree, with a great deal of difficulty—hence the term, frustration level. The 95% confidence interval for this is 37% ± 8.11 which places the upper and lower limits on the outside perimeters of frustration-level reading comprehension.

The second treatment group (CLARITY=80) had a 95% confidence interval of 53.86% ± 6.89; the mean for this group fell within the range established by Bormuth for instructional level reading comprehension (44% - 57%). Instructional level reading performance has been defined as a reader's ability to comprehend the reading material, with the help of an instructor.

The third treatment group, in which "CLARITY" equaled 86, had a 95% confidence interval of 51.00% ± 7.43. This too, placed those scores within the instructional reading level. The means from the fifth and sixth treatment groups fell at the junction between the instructional and the independent reading levels. Performance of students scoring within the range established for the independent level indicates that they are able to read the material without the aid of an instructor—with little difficulty.
The fifth treatment group had a 95% confidence interval of 58.33\% \pm 7.79, while the sixth treatment group had a 95% confidence interval of 57.25\% \pm 10.08. The last treatment group, in which only 6% of the passage was nonsense, had a 95% confidence interval of 55.80\% \pm 9.33 (see Figure 3 below).

Figure 3
95% Confidence Intervals
2. \( H_0 \): No comprehension will occur in the control group which receives the scrambled text (CIARITY = 0).

The lower control group (CIARITY = 0) had a 95% confidence interval of \( 53.12\% \pm 3.88 \) (see Table 7). The null hypothesis is, therefore, rejected. Not only did the mean of this group extend into the instructional level according to Bormuth's performance levels, but even with establishing a 95% confidence interval, the upper and lower limits still remained within the perimeters of the instructional level.

In looking at Table 7, it is obvious that something different is happening with this control group as compared to the other eight groups. When these data (all nine groups) were submitted to regression analysis using orthogonal polynomials, a quadratic component emerged in addition to the linear one. When the control group (CIARITY = 0) was taken out, however, a linear equation became the best characterization of the means. Had there been a quadratic fit to these data in addition to a linear component, however small, it would have emerged in several of the statistics. This, however, was not the case as seen from a discussion of the following statistics.
The F ratio for the quadratic equation was not significant \( (F = 1.884) \), nor was there significant change in the \( R^2 \) from the linear to the quadratic fit: \( .01162 \).

This led the researcher to conclude that an aspect was being tested with the first control group other than the aspect intended: namely, either the subjects' problem-solving ability or readability of the text but not, in this case, comprehension of the control text. This will be discussed in detail in Chapter V.

3. \( H_0 \): No comprehension will occur in the control group which receives the passage with the straight copy of the text (CIARITY = 100).

The second control group (CIARITY = 100) had a 95% confidence interval of \( 60.22\% \pm 10.66 \) (see Table 7). The mean for this particular group in this study clearly places the students' performance into the independent reading comprehension level as established by Bormuth (1968). The null hypothesis for this control group is, therefore, rejected.

**Statistics Indicating a Linear Trend**

A step-wise regression was performed on these data by using the "PCT" scored on the criterion measure to find an equation of the best fit for the means. The results
Table 8
Analysis of Variance and Orthogonal Polynomial Regression Statistics
Breakdown for Step 1

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>PCT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables entered on Step 1</strong></td>
<td><strong>CIARITY</strong></td>
</tr>
<tr>
<td>Multiple R</td>
<td>0.49531</td>
</tr>
<tr>
<td>R Square</td>
<td>0.24533</td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.23513</td>
</tr>
<tr>
<td>Standard Error</td>
<td>11.57825</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables in the equation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td>CIARITY</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>
of step 1 are presented in Table 8. The F ratio is 24.06 (p < .01) which indicates that the linear effect of clarity accounts for a significant amount of variance. In other words, clarity has a linear effect upon comprehension, according to the results of this study. Because there is only a single variable at this step, multiple R is really a simple r; this represents the correlation coefficient of observed to predicted scores (r = 0.49531). The R² for the linear equation is 0.24533 which shows the amount of "PCT" which can be explained by the independent variable "CLARITY". The standard error shows the variability of the scores around the predicted linear regression line, expressed in standard units.

Column B in Table 8 contains the values of the slope and intercept of the linear equation for these data:

\[
PCT = 0.6818 \times \text{CLARITY} + (-)6.3000
\]

This equation determines the linear regression fit: it states that if a subject receives the control text in which "CLARITY" = 0, then his/her predicted score would be -6.3%. If a student receives a normal passage in which no permutations, additions, or deletions have occurred, his/her predicted score would be 61.88%.
A step-wise regression analysis was performed; the results are presented in Table 9. This table breaks down and summarizes the significance of the linear equation, as opposed to any other, for the results. As can be seen from Table 9, $R^2$ in the linear equation is .24533 and in the quadratic .25695. The difference between the linear and quadratic equation is the $R^2$ change; in this case the difference is .01162. This small difference between the linear and quadratic fit indicates that no significant quadratic relationship exists between "PCT" and "CLARITY."

Results Obtained by Means of Scheffe

Although 95% confidence intervals had been previously established with these data, it was decided that a more stringent and conservative estimation of the means was needed. The Scheffe test, a conservative measure, was used to determine if any of the means from the subpopulations were significantly different. This test was employed for the following reasons:

1. The means from each of the subpopulations only represent a point estimation of a theoretical mean within the given range of the 95% confidence interval; this interval was previously determined,
Table 9

Step-wise Regression Summary Table

<table>
<thead>
<tr>
<th>Variable</th>
<th>Multiple R</th>
<th>$R^2$</th>
<th>$R^2$ Change</th>
<th>Simple R</th>
<th>B</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td>0.49531</td>
<td>0.24533</td>
<td>0.24533</td>
<td>0.49531</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.50690</td>
<td>0.25695</td>
<td>0.01162</td>
<td>0.48761</td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-6.300</td>
<td></td>
</tr>
</tbody>
</table>
however, by using the results gained from the data and, therefore, more dependent upon the variance of the data.

2. The test acts as a further safeguard to balance the small N.

3. The test is also an assurance that Type I error will be held constant at .05.

4. Because the test is conservative, it protects against all possible comparisons of any two means.

The Scheffé found that there was a significant difference between treatment 1 and treatment 5; treatment 1 and treatment 6; and, treatment 1 and control 2.

Summary

Through submission of the data to trend analysis using orthogonal polynomials, the results indicate the rate at which comprehension changes as the rate of passage difficulty changes:

\[ PCT = 0.6818 \times CLARITY + (-) 6.300 \]

The results also indicate that a linear equation is the best possible characterization of the means (see Figure 4).
Figure 4

Illustration of the Slope and Intercept of Line
Chapter V
Discussion

Introduction
This chapter will discuss the results of the experiment in their relation to the research questions posed in Chapter I. In addition, the theoretical bases of the study will be reviewed in light of the findings from the present study, as well as a discussion concerning the elimination of the first control group from the analysis. Finally, the discrepancies between the pilot and the research study will be discussed as they relate to the particular findings of the experiment as well as to current knowledge and theories in the fields of psycholinguistics and second language learning.

Research Questions and the Results
Chapter I proposed that many questions relating to the role of new-word density, and its effect upon reading comprehension, needed answers. The research questions
Table 10

Cloze Test Criterion Scores Equivalent to Traditionally Accepted Performance Levels

<table>
<thead>
<tr>
<th>Traditional Scores</th>
<th>Bormuth</th>
<th>Rankin</th>
<th>Anderson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional</td>
<td>75%</td>
<td>45%</td>
<td>41%</td>
</tr>
<tr>
<td>Independent</td>
<td>90%</td>
<td>52%</td>
<td>61%</td>
</tr>
<tr>
<td>Frustration</td>
<td>--</td>
<td>35%</td>
<td>--</td>
</tr>
</tbody>
</table>
asked what density of new-words (ratios) could be established that could be said to interfere with comprehension. This question was further expanded to include the notion of performance levels established by Bormuth's (1968) study (frustration, instructional, and independent). What was specifically being asked, then, was whether ratios of new-word densities can be established for each of the three performance levels of reading comprehension established by Bormuth in his 1968 study? A review of the cloze test criterion scores and their equivalence to traditionally held performance levels can be found in Table 10.

According to Bormuth's study, students scoring between 35% - 44% on a cloze test for comprehension were said to be reading at the frustration level. The results of this study place the limits of the first treatment group within that range. The first treatment group, in which "CLARITY" = 67, had a 95% confidence interval of 37% ± 8. This confidence interval gives an assurance that 95 out of 100 times the means of the samples tested will fall within the limits described (28.86% - 45.11%). The lower limit falls somewhat below the lower boundary instituted by Bormuth for frustration level; however,
the upper boundary is within the prescribed limits as well as the mean which, for this particular sample, was 37%.

This finding presents new evidence on the level of vocabulary density that can be tolerated by an English as a Second Language student while reading.

Prior research by Holley (1973) had indicated that American students learning German tolerated and comprehended one new word for every 15 known words. Her research did not, however, deal with new-word densities below the ratio of one-to-fifteen; nor did it compare its results to pre-established, performance levels for reading comprehension. MacGinitie (1963) discovered an optimal point for "cloze" by deleting varying proportions of words in context. He surmised from the results of his research that deletions which leave a context of less than four words are very difficult to reconstruct; this same experiment lent credence to his hypothesis that every fifth word deletion was optimal. In the present study, when a total of one-third of the English passage was deleted and nonsense syllables were inserted in their place, the English as a Second Language students were able to reconstruct the cloze criterion measure at the frustration level ($\bar{x} = 37\% \pm 8.11$); this study, then, supports MacGinitie's
hypothesis and established a one-third ratio of new-to-known words as the limit tolerated for some reading comprehension to occur. The following is an illustration of the groups with their differing ratios of nonsense syllables (see Figure 5).

The research questions also asked if a ratio of new-word density could be established for instructional level and independent-level reading comprehension to take place; these questions are somewhat harder to answer with the same definitiveness. This is because of the statistics involved in the analysis: in this study, the 95% confidence intervals established that treatment 1 was different from treatments 2, 4, 5, 6, 7 and control 2; these intervals further established that treatments 2 through 7 were not different from each other or from control 2. Further analysis by means of the Scheffé test found significant differences only between treatments 1 and 5; treatments 1 and 6; and treatment 1 and control 2. Taking these two different types of analyses into account allows the establishment of frustration-level reading performance; but it does not, strictly speaking, define a single ratio for the beginning of instructional level, the ending of instructional level, or the beginning of
Diagram of Nonsense Words Inserted According to Ratio

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Pattern on Criterion Measure</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:2</td>
<td>--x--x--x--x--x--x--x--x--x--x--x--x</td>
<td>37.00</td>
</tr>
<tr>
<td>1:4</td>
<td>----x------x------x------x------x------x------x</td>
<td>53.85</td>
</tr>
<tr>
<td>1:6</td>
<td>---------x------------x------------x------------x------------x</td>
<td>51.00</td>
</tr>
<tr>
<td>1:8</td>
<td>-----------x----------x----------x----------x----------x</td>
<td>55.22</td>
</tr>
<tr>
<td>1:10</td>
<td>----------------x-------------------x-------------------x</td>
<td>58.33</td>
</tr>
<tr>
<td>1:12</td>
<td>-------------------x-------------------x-------------------x</td>
<td>57.25</td>
</tr>
<tr>
<td>1:14</td>
<td>---------------------x---------------------x---------------------x</td>
<td>55.80</td>
</tr>
</tbody>
</table>

Figure 5
independent-level reading performance.

Looking at the means only, and interpreting them less strictly, it is possible to make certain conclusions about the groups under study. Treatment 2 (CLARITY = 80) had a 95% confidence interval of 46.96% - 60.74% ($\bar{X} = 53.85\%$). This indicates that with 20% of the passage missing, that these subjects were able to read and understand at the instructional level. Likewise, treatment group 3 (CLARITY = 86) had a 95% confidence interval of 43.57% - 58.42% ($\bar{X} = 51\%$). This treatment falls within the range established by Bormuth for instructional level as does the next group in which CLARITY = 89; the 95% confidence interval established for this group (4) was 52.01% - 58.43% ($\bar{X} = 55.22\%$). It could be defended, therefore, that when 80%, 86%, and 89% of intermediate-level material is known, that the intermediate English as a Second Language student will be able to read at the instructional level; that is, with the aid of an instructor.

The last three treatments and second control could be said to be mediating between the instructional level and the independent level of reading comprehension. Treatment group 5 in which 91% of the text remained had a 95% confidence interval of 50.54% - 66.12% ($\bar{X} = 58.33\%$);
group 6 in which 93% of the text remained had a 95% confidence interval of 46.75% - 67.74% (\(\bar{x} = 57.25\%\)); group 7 in which 96% of the text remained, had a 95% confidence interval of 46.46% - 65.13% (\(\bar{x} = 55.80\%\)); and the second control, in which all of the words were regular English had a 95% confidence interval of 49.55% - 70.88% (\(\bar{x} = 60.22\%\)). These last groups had only between 9% and 0% of the text missing; in addition, their means ranged from a low of 55% to a high of 60%. This range clearly borders the 57% and above established by Bormuth for independent-level reading comprehension performance, and therefore, could be defended thus: with as much as 9% of the material unknown, the intermediate-level English as a Second Language student will still be able to comprehend the intermediate-level English as a Second Language material at the high end of the instructional level.

The last research question asked what effect, if any, increasing the ratio on nonsense syllables on measures of reading comprehension has on intermediate-level English as a Second Language students. When English as a Second Language students were asked to read selections with varying proportions of nonsense syllables inserted in the selection to replace regular English words, it was found that increasing these ratios had a linear effect upon comprehension. The slope and intercept was defined by the following equation:
PCT = \(0.6817 \times CLARITY + (-) 6.3\)

This means that for every 1% increase in score, "CLARITY" will increase by 0.68. Although this study only dealt with the following ratios 1:2 (67), 1:4 (80), 1:6 (86), 1:8 (89), 1:10 (91), 1:12 (93), 1:14 (95), and 100—by using this equation, any other ratio can now be determined as in Table II.

**Theoretical Bases Reviewed in Light of Research Findings**

It was stated in Chapters I and II that the theoretical underpinning for this study was the psycholinguistic reading model. Psycholinguistics offers a theory about how the written word is comprehended, irrespective of language background. There are basically three central themes to psycholinguistic thought concerning the reading process:

1. Only a small part of the information necessary for reading comprehension comes from the written page.
2. Comprehension can precede the identification of individual words.
3. Fluent reading is not decoding to spoken language.
<table>
<thead>
<tr>
<th>% of Known Words</th>
<th>Expected Score on Cloze</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% CIARITY</td>
<td>-6.30%</td>
</tr>
<tr>
<td>10% CIARITY</td>
<td>0.52%</td>
</tr>
<tr>
<td>15% CIARITY</td>
<td>7.33%</td>
</tr>
<tr>
<td>20% CIARITY</td>
<td>10.77%</td>
</tr>
<tr>
<td>25% CIARITY</td>
<td>14.15%</td>
</tr>
<tr>
<td>30% CIARITY</td>
<td>17.56%</td>
</tr>
<tr>
<td>35% CIARITY</td>
<td>20.97%</td>
</tr>
<tr>
<td>40% CIARITY</td>
<td>24.38%</td>
</tr>
<tr>
<td>50% CIARITY</td>
<td>27.88%</td>
</tr>
<tr>
<td>55% CIARITY</td>
<td>31.19%</td>
</tr>
<tr>
<td>60% CIARITY</td>
<td>34.60%</td>
</tr>
<tr>
<td>65% CIARITY</td>
<td>38.07%</td>
</tr>
<tr>
<td>70% CIARITY</td>
<td>41.42%</td>
</tr>
<tr>
<td>75% CIARITY</td>
<td>44.83%</td>
</tr>
<tr>
<td>80% CIARITY</td>
<td>48.24%</td>
</tr>
<tr>
<td>85% CIARITY</td>
<td>51.64%</td>
</tr>
<tr>
<td>90% CIARITY</td>
<td>55.05%</td>
</tr>
<tr>
<td>95% CIARITY</td>
<td>58.46%</td>
</tr>
<tr>
<td>100% CIARITY</td>
<td>61.87%</td>
</tr>
</tbody>
</table>
Clearly, these statements indicate that there must be a process which characterizes and describes the psycholinguistic reading model. A description of this model helps explain the process: the psycholinguistic reading model is like an information processing system in which visual information is picked up by the eye after having been directed to do so by the brain. As the information begins flowing in, the brain processes it for meaning using various methods such as the knowledge by the speaker of the syntax of the language, the speaker's knowledge of the world and, in conjunction with this, the content of the information. Reasonably, the more highly developed these two attributes are, the more quickly and accurately the brain will be able to process the information. This processing ability relies on three internal cognitive cueing systems in order to allow the brain to make predictions and confirm or disconfirm hypotheses about the reading passage.

The results of this study indicate that some kinds of processes were at work while the subjects were reading; in other words, it is clear that some kind of strategies were being employed in order to reduce uncertainty and derive meaning from the reading passage. One has only to
look at the varying amounts of information remaining in the selections in order to realize that meaning was derived more through the use of skills involving knowledge of the grammar and knowledge of the world than through the decoding of the visual marks on the page. A further comparison of the means, correlated with the percentage of actual words remaining in the text, underscore this point (see Figure 5).

Stephens (1977) believes that students proficient in their native language have the essential skills to recognize the important linguistic factors necessary to derive meaning from the reading selection in the target language. His work and that of other psycholinguists have been based upon the assumption that all natural languages share certain characteristics: one such characteristic is the ability to relate a semantically interpretable deep structure to a phonetically interpretable surface structure through the syntax of the language. Thus, once the fundamentals of the reading process have developed in the native language, the ability to transfer that knowledge to the second language would develop without having to understand the fundamentals of the process with each new language. The results of
this study seem to underscore this assumption.

The subjects in this experiment were intermediate-level English as a Second Language students at the University of Cincinnati. Most of them were graduate students; most were in the physical sciences; some, of course, were undergraduates. Their ages ranged from 17 to 30 but the common denominator was that they were all classified as intermediate-level English as a Second Language students. The selection, "The Autobiography of Abe Lincoln," was carefully chosen: not only did it represent certain essentials that characterize the intermediate-level English as a Second Language reading text, but also, though while relating the history of the life of a famous American, it was, nevertheless, information not commonly known and, hence, not the kind of information that could be used to answer the cloze comprehension test without understanding the story (see Appendix G).

Stephen's assumption that reading skills are transferable from native language to second language is reinforced by the results obtained from this particular sub-population. The first treatment group received a passage in which 33% of the total passage was missing; the mean for this particular sample was 37% which indi-
cates that some comprehension of the passage took place. The likelihood of anyone reading a selection in which a total of 33% was missing and still being able to comprehend the material is not great unless some strategies to derive meaning are employed. It is in this process that psycholinguists are interested; for them, reading is not a simple mechanical skill nor a completely visual process—comprehension is a thought process which one goes through while reading, in order to derive meaning. It is the assumption of the researcher that these same kinds of processes were being employed by the subjects in this research experiment.

Control group 1 gives a further indication that prediction strategies were being used by the subjects. This group received the selection in which every sentence was completely scrambled. A 95% confidence interval was established for this group of $49.25\% - 57\% (\bar{x} = 53.12\%)$, a standard deviation of only 4.64. When the results of this group were submitted with the results from the other groups in the experiment, it was found that there was a quadratic component in addition to the linear. It was further found, however, that the F ratio for the quadratic equation was not significant; nor was there a significant
change in the $R^2$ from the linear to quadratic fit. Had either one of these statistics been significant, it could have been concluded that there was, indeed, a quadratic component in addition to the linear. This was not, however, the case. It could only be concluded, therefore, that some other aspect of reading was being tested, but not comprehension of the selection.

It is the feeling of the researchers that what was being tested by this group was their collective ability to use predictive strategies to derive meaning from the reading selection, or their ability to solve problems. This group received the selection in which all of the sentences were scrambled; the text was, however, all English words, which it is important to note. It is important because it is highly likely that these students used their knowledge of the grammar language as well as their conscious knowledge of the statistics or redundancies of English, in addition to their general knowledge of the world, in order to resolve the uncertainty and to derive meaning from the selection.

Further, it is extremely interesting to note the small differences in the means between control 1 and treatment 2. Control 1 for this particular sample had a mean of 53.12% while the mean for treatment 2 was 53.85%. If the subjects assigned to the control group were able to piece together enough information about the
story, then they would have had some chance of accurately predicting the cloze blanks on the comprehension test. In addition, this group would not have had the interference or large amounts of nonsense syllables to distract them as did treatment group 1; it is hypothesized, therefore, that by using accurate prediction strategies, this group should have obtained a mean score somewhere between the means of treatment 1 and treatment 2. This is precisely what happened. Possession of fairly accurate prediction strategies enabled them to recognize the basic linguistic factors in each sentence. When they took the cloze test of comprehension in which every fifth word had been deleted, they were able to put the basic linguistic factors derived from the story into their proper perspective, and thus correctly fill the cloze blanks.

Another possible explanation for the high scores in control group 1 is that these subjects used the cloze test not as a test for comprehension, but as a test of readability of the material. In that case it would explain the similar results between control 1 and treatment group 2; it would also support MacGinitie's and Bormuth's claims that every fifth word deletion was optimal for discovering either readability or comprehension of a reading selection. As a test of readability of the material, the means of the subjects for instructional level should fall within the range of 44% - 57%; a 95% confidence interval was established.
for this group of 53.12±3.88, which clearly puts this group within the instructional range. It must be remembered, however, that the cloze test for readability as well as for comprehension is based upon psycholinguistic assumptions about the learner and the language: namely, how the subject will interact with the reading situation. Making the case that the subjects in control group 1 were testing the readability of the passage rather than comprehension of the scrambled text, only further intensifies the claim that psycholinguistic-type strategies were being employed. In examining all the evidence, it becomes increasingly apparent, however, that readability can only partially explain the results.

Using the hypothesis that it was readability of the passage which was actually tested in control group 1, to account for the divergent scores, does not, however, explain the startling amount of context recorded in their answers, nor the similarity in their scores with treatment 2. It can only be concluded that these intermediate-level English as a Second Language students have become expert at solving language problems. This expertise at problem solving can only be based upon certain kinds of knowledge about the English language and language in general; this includes:

1. knowledge of the statistics or redundancies of the language.
2. knowledge of English syntax

3. knowledge of certain language universals (i.e. surface structure differs from language to language, but all natural languages have some kind of order imposed upon their surface structure)

4. That this surface structure merely reflects and allows the reader to interpret a semantically-based deep structure.

It is interesting to speculate, based upon the above information, that the students in control group 1 attacked the scrambled passage much in the same way that they attack all written English and, furthermore, that since these students came from areas in which their native languages have widely divergent surface structures from English, it is possible that the scrambled passage offered an identical or very similar puzzle to the one they initially had had to solve in order to begin reading English. Further substantiation of the psycholinguistic reading process will be discussed in the next section.

Discrepancies between the Pilot and Experimental Studies

Although both the pilot and the experimental studies resulted in a linear trend as the best possible fit for the data, it is obvious from a comparison of the means in each study that there are some fundamental differences, attributed by the researcher as reflecting basic underlying differences between the second-dialect and second-language students
The developmental students used in the pilot were, for the most part, second-dialect speakers. These American students were enrolled in a developmental English/reading program in order to develop their English linguistic skills. Developmental students typically share the following characteristics:

1. inability to read fluently in native language
2. lack of a firm grasp of English syntax which leads to an
3. unawareness of statistics of language and lack of sensitivity to language
4. inability to recognize words in terms of linguistic factors necessary to derive meaning in reading
5. lack of experience in manipulating language; tendancy to have limited vocabulary
6. in many cases, lack of motivation to develop linguistic skills
7. lack of general knowledge of the world.

These characteristics help to explain the differences in the results between the first control group in each study. The mean for control group 1 in the pilot study was 23.71% while the mean for control group 1 in the experimental study was 53.21%. These results indicate that variables other than the effect of vocabulary density upon comprehension was being tapped by the research design. With the develop-
Figure 6

Comparison of Means between Pilot and Experimental Studies
mental students, there were but infrequent attempts at trying to read the scrambled passage; because the subjects had not attempted to read the selection, they verbally indicated that they could not do the comprehension test. With only two exceptions, all scores fell well below 29%. The foreign students, on the other hand, had a mean of 53.12% and a standard deviation of 4.6, which indicates the small amount of variability in the scores. The differences in the results between these two groups indicate that such variables as experience with manipulating language, ability to recognize words in terms of linguistic factors, awareness of language redundancies, knowledge of syntax, knowledge of the world, and motivation are the factors which are being tapped by this control passage.

These factors are common to the English as a Second Language student in this country, but not to the second-dialect student. Moreover, these factors represent, in general, considerations important in explaining the psycholinguistic reading model. It seems therefore reasonable to assume that if the second-language student is well versed in, for instance, syntax, and the second-dialect student is not, the second-language student's ability to make accurate predictions will far exceed that of those students whose grasp of syntax is vague and uncertain. The conclusion can be drawn and defended, therefore, that the psycholinguistic reading model with its emphasis upon strategies
for prediction explains the resulting differences from these two studies. The results further led the researcher to conclude that while Labov may be very right in assuming shared linguistic characteristics between second-dialect speakers and second-language speakers, that these characteristics are certainly more sociolinguistic than psycholinguistic in nature.

Summary

Results from the experimental study indicate that the tolerance of unfamiliar words in a reading selection is much higher for English as a Second Language students in this country than had previously been thought possible. It was found that for this particular sample even as high a density as one new word for every two familiar words resulted in a frustration-level reading comprehension; when the density was diluted to one new word for every four known words, the reading performance of the students jumped to the instructional level. At that point the subjects could be said to be mediating in performance between the instructional and independent reading levels. When a comparison was made between the performance of the developmental American students for which English is a second dialect, and the performance of the English as a Second Language students in the first control group, it was proposed and defended that the basis for the more fluent reading comprehension scores by the English as a Second Language
students could be understood by the theories of the psycho-
linguistic reading model. It was further argued that it
is a system of prediction and strategies founded upon
linguistic and other knowledge that enabled the English
as a Second Language student to outdistance by far the
developmental student in performance.
Chapter VI

Conclusions, Summary & Overview, and Implications

Conclusions

In a university setting, fluent reading is one of the most basic academic skills needed for survival. It is often assumed that the university level English as a Second Language student possesses a higher level knowledge about reading English texts. Sadly, this is often not the case. Lack of knowledge about reading is all too apparent from a brief survey of any university English as a Second Language class—asked what gives them the most difficulty, the recurring complaint is usually too little time and too much reading.

What differentiates a fluent reader from one who ploddingly labors over a text with the assignment in one hand and a dictionary in the other? From the results of the study, it is felt that the fluent reader regularly employs predictive strategies. It is these strategies, rather than the possession of a large vocabulary in English, that denote success.

Educators in the field of second language acquisition have long assumed that comprehension is adversely affected by
the introduction of too many unfamiliar words, i.e. new-word density. New-word density, according to the results of this study, is not responsible for the drudgery of long reading assignments; the important factor seems to be the utilization of essential knowledge about the language in making accurate predictions.

Summary and Overview of Experiment

Design. This study employed an experimental research design. The experiment consisted of randomly assigning students to one of the two control or seven treatment groups. The first control group read a passage in which every sentence had been scrambled; the second control group read a straight rendition of the text. Treatment group one read a passage in which every third word was deleted and a nonsense syllable inserted in its place: thus establishing a ratio of 1:2 for the purposes of statistical analysis. The second treatment group had a reading passage with a ratio of 1:4 which meant for every four English words, one nonsense syllable was inserted. The ratios 1:6, 1:8, 1:10, 1:12, and 1:14 comprised the remaining five treatments. The reading was the same in each group; the variable that differentiated one group from another was the frequency of nonsense syllables inserted into the mutilated text. After reading the passage, each student was required to fill in the
same cloze test of comprehension over the reading selection (see Table 6, p. 73).

Results. The results were scored by tallying only exact word replacement; each correctly closed blank was worth one point. The results from each group were then submitted to regression analysis using orthogonal polynomials. The analysis determined that a linear fit was the best characterization of the means; it also yielded a linear equation to be used in predicting future cloze test scores. The means were then compared to performance levels previously established by Bormuth (1968) for the cloze test (see Table 1, p. 38). This comparison allowed the researcher to judge how effectively an English as a Second Language student could tolerate the differing high frequencies of new words. The following table illustrates the results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td>0</td>
<td>53.12</td>
</tr>
<tr>
<td>Clarity</td>
<td>67</td>
<td>37.00</td>
</tr>
<tr>
<td>Clarity</td>
<td>80</td>
<td>53.85</td>
</tr>
<tr>
<td>Clarity</td>
<td>86</td>
<td>51.00</td>
</tr>
<tr>
<td>Clarity</td>
<td>89</td>
<td>55.22</td>
</tr>
<tr>
<td>Clarity</td>
<td>91</td>
<td>58.33</td>
</tr>
<tr>
<td>Clarity</td>
<td>93</td>
<td>57.25</td>
</tr>
<tr>
<td>Clarity</td>
<td>94</td>
<td>55.80</td>
</tr>
<tr>
<td>Clarity</td>
<td>100</td>
<td>60.22</td>
</tr>
</tbody>
</table>
Looking at the means only, and interpreting them less strictly than with a conservative analysis like the Scheffe-test, it is possible to make certain conclusions about the groups under study. As is readily apparent from glancing at the means, each group tolerated with varying degrees of comprehension a fairly high new-word density. If new-word density was solely responsible for labored reading, these results would have been different. It seems plausible, therefore, that some other factor is at work.

Psycholinguistics offers a theory regarding how the written word is comprehended, irrespective of language background. There are basically three central themes to psycholinguistic thought concerning the reading process:

1. Only a small part of the information necessary for reading comprehension comes from the written page.
2. Comprehension can precede the identification of individual words.
3. Fluent reading is not decoding to spoken language.

A description of this model help explain the process. According to the psycholinguistic reading model, reading can be viewed as an information processing system in which visual information is picked up by the eye after having been directed to do so by the brain. As the information begins flowing in,
the brain processed it for meaning using various methods such as the knowledge by the speaker of the syntax of the language, the speaker's knowledge of the world and, in conjunction with this, the content of the information. Reasonably, the more highly developed these attributes are, the more quickly and accurately the brain will be able to process the information. The processing ability relies on these internal cognitive cueing systems which allows the brain to make predictions and confirm or disconfirm hypotheses about the reading passage.

The results of this study clearly indicate that some kinds of strategies were being employed in order to reduce uncertainty and derive meaning from the reading passage. One has only to look at the varying amounts of information remaining in the selections in order to realize that meaning was derived more through the use of skills involving knowledge of the grammar and knowledge of the world than through the decoding of the visual marks on the page.

Further, Stephens (1977) wrote that students proficient in their native language have the essential skills to recognize the important linguistic factors necessary to derive meaning from the reading selection in the target language. His work and that of other psycholinguists, psychologists, and linguists have been based upon the assumption that all natural
languages share certain characteristics: one such characteristic is the ability to relate a semantically interpretable deep structure to a phonetically interpretable surface structure through the syntax of the language. Thus, once the fundamentals of the reading process have developed in the native language, the ability to transfer that knowledge to reading a second language would develop without having to understand the fundamentals of the process with each new language. The results of this study seem to support this assumption.

Implications

Since expertise at language problem solving is based upon certain kinds of knowledge about the English language, it would seem that an effort should be made to make the English as a Second Language student conscious that he possesses this knowledge. Once aware, his predictive strategies and hence his reading fluency will be increased. This knowledge would include:

1. knowledge of the statistics or redundancies of the language
2. knowledge of English syntax
3. knowledge of certain language universals and, hence, transferrable knowledge from the first to the target language reading process; e.g., knowledge that language
surface structure differs from language to language but all natural languages have some kind of order imposed on their surface structure.

4. knowledge that the surface structure merely reflects and allows the reader to interpret a semantically based deep structure; such an understanding is based on a knowledge of semantic properties of sentences that are obscured in surface structure. This includes:

a. synonymy resulting from the deletion of indefinite elements
b. synonymy resulting from identity deletions
c. ambiguity
d. generative ability
e. recognition of ungrammaticality

In conclusion, it would appear that the needs of university-level English as a Second Language students would be best served by teaching them how to use their knowledge of the essentials of the English language in order to make accurate predictions.

Recommendations and Limitations

The subjects in this experiment were intermediate-level English as a Second Language students at the University of Cincinnati. Most of them were graduate students; most were in the physical sciences; some, of course, were undergraduates.
Their ages ranged from 17 to 30 but the common denominator was that they were all classified as intermediate-level English as a Second Language students. It is possible to conjecture that the conclusions reached through an analysis of this study would be applicable to more advanced English as a Second Language university students; it is important, however, that this study be replicated with more advanced students to judge whether this assumption is justified. On the other hand, the conclusions from this study are not likely to hold true for lower-level English as a Second Language students, mainly because these students lack the essential knowledge about English to make accurate predications. This assumption, however, also needs to be tested. In addition to replicating this research with other university-level students, the researcher feels it would be interesting to see if the results would remain the same with adults learning English in adult basic education programs. Finally, it is important to note that the results from this study cannot be generalized to any English as a Second Language population other than intermediate-level university students in this country.
BIBLIOGRAPHY


Archer, E. J. "A re-evaluation of the meaningfulness of all possible CVC trigrams." Psychological Monographs, 74 (1960), 497.


-----. "Comparisons among cloze test scoring methods." In J. A. Figurel (Ed.), Reading and Inquiry. 10 (1964), 283-86.


Gibson, S., Yonas, M. "Utilization of spelling patterns by deaf." In Levin & Williams (Eds.), Basic studies on reading. 1970, 54-73.


------. "Do you have to be smart to read? Do you have to read to be smart?" Reading Teacher, 28 (1975), 625-32.


Henderson, E. H. "The role of skills in teaching reading." In Theory Into Practice, 16 No. 5 (1977), 348-56.


Noble, C. "Measurements of association value (a), rated associations (a'), and scaled meaningfulness (m') for the 2100 CVC combinations of the English alphabet." Psychological Reports, 8 (1961), 487-521.

Nodine, C. F., Hardt, J. V. "A measure of pronunciability of CVC trigrams." Behavior research methods and instrumentation, 1 No. 6 (1969), 210-16.


"Cloze, discourse, and approximations to English." (EDRS: ED 107 144) 1975.

"Scoring methods and difficulty levels for cloze tests of proficiency in English as a second language." Modern Language Journal, 56 No. 3 (1972), 151-58.


Pavlak, S. A. "Excerpts from significant research on comprehension (1948-1972)." (EDRS: ED 095 506) 1974.


Richards, J. C. "The role of vocabulary teaching." TESOL Quarterly, 10 No. 1 (1976), 77-89.


Twaddle, F. "Vocabulary expansion in the TESOL classroom." TESOL Quarterly, 7 No. 1 (1973), 61-78.


Appendix A

Instructions for Reading Selection

Introduction

Nothing is more frustrating for foreign students in this country than to be required to read English language material that is difficult because of unfamiliar vocabulary. In order to solve this problem, you have been selected to participate in an education research study designed to provide English as a Second Language instructors with empirical data that will enable them to accurately choose appropriate reading selections for students.

Instructions

Before you is a reading selection in which English words have been deleted and nonsense syllables inserted in their place. You will readily be able to tell which words are nonsense as they will not look like English and will all have just (3) three letters. You are to read the text as best you can; the object of the reading is to try and understand the story. There is no time limit so you may re-read the story if you wish, but once you finish you will not be allowed to look back at the story to answer the comprehension questions. When you have finished the reading and are ready to take the comprehension test, please raise your hand and someone will collect your reading selection and give you the questions over the passage.
Appendix B

Instructions for Cloze asure

______________________________
name

______________________________
department/year

Instructions

The attached test has been made by simply deleting every fifth word from the entire story that you just read. Your task is to first read through the story and then to fill in each blank with the English word that makes the most sense to you. You are to write only one word in each blank. If you are not sure of the word, do not be afraid to guess. All of the blanks can be answered with ordinary English words. Wrong spelling will not be counted incorrect if the test grader can tell what word you intended. The test is not timed but you should be able to finish in about 20 minutes. When you finish please raise your hand and the test will be collected. If you have any questions, please feel free to ask. Many thanks!

*The results of this test will be held in strictest confidence. It is important to have your name so that this score can be correlated with the Michigan Test of English Language Proficiency and/or TOEFL.*
Appendix C

Fry's Readability Graph

Figure 2

GRAPH FOR ESTIMATING READABILITY—EXTENDED

By Edward Fry, Rutgers University, Reading Center, New Brunswick, N.J. 08904

Expanded Directions for Working Readability Graph

1. Randomly select three (3) sample passages and count out exactly 100 words each, beginning with the beginning of a sentence. Do count proper nouns, initializations, and numerals.

2. Count the number of sentences in the hundred words, estimating length of the fraction of the last sentence to the nearest one-tenth.

3. Count the total number of syllables in the 100-word passage. If you don't have a hand counter available, an easy way is to simply put a mark above every syllable over one in each word, then when you get to the end of the passage, count the number of marks and add 100. Small calculators can also be used as counters by pushing numeral 1, then push the + sign for each word or syllable when counting.

4. Enter graph with average sentence length and average number of syllables; plot dot where the two lines intersect. Area where dot is plotted will give you the approximate grade level.

5. If a great deal of variability is found in syllable count or sentence count, putting more samples into the average is desirable.

6. A word is defined as a group of symbols with a space on either side; thus, Joe, IRA, 1945, and & are each one word.

7. A syllable is defined as a phonetic syllable. Generally, there are as many syllables as vowel sounds. For example, stopped is one syllable and wanted is two syllables. When counting syllables for numerals and initializations, count one syllable for each symbol. For example, 1945 is four syllables, IRA is three syllables, and & is one syllable.

Note: This "extended graph" does not oust or render the earlier (1963) version inoperative or inaccurate; it is an extension. (REPRODUCTION PERMITTED—NO COPYRIGHT)
Appendix D

Description of Subpopulations

Pilot

<table>
<thead>
<tr>
<th>VARIABLE FOR ENTIRE POPULATION</th>
<th>CODE</th>
<th>VALUE LABEL</th>
<th>SUM</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>VARIANCE</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.</td>
<td></td>
<td>2558.0000</td>
<td>38.1791</td>
<td>15.2011</td>
<td>233.5129</td>
<td>67</td>
</tr>
<tr>
<td>CLARITY</td>
<td>67.</td>
<td></td>
<td>166.0000</td>
<td>23.7143</td>
<td>15.7450</td>
<td>247.5048</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>80.</td>
<td></td>
<td>221.0000</td>
<td>31.5714</td>
<td>15.8204</td>
<td>250.2870</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>85.</td>
<td></td>
<td>282.0000</td>
<td>35.2260</td>
<td>15.5632</td>
<td>242.2143</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>89.</td>
<td></td>
<td>340.0000</td>
<td>43.2500</td>
<td>14.2503</td>
<td>203.6714</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>91.</td>
<td></td>
<td>304.6667</td>
<td>40.5714</td>
<td>17.0182</td>
<td>284.6190</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>94.</td>
<td></td>
<td>229.0000</td>
<td>32.7457</td>
<td>13.7215</td>
<td>189.7160</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>99.</td>
<td></td>
<td>255.0000</td>
<td>32.7457</td>
<td>10.7038</td>
<td>114.5714</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>100.</td>
<td></td>
<td>324.0000</td>
<td>40.5714</td>
<td>14.1328</td>
<td>200.2657</td>
<td>7</td>
</tr>
<tr>
<td>CLARITY</td>
<td>100.</td>
<td></td>
<td>314.0000</td>
<td>44.8971</td>
<td>7.8619</td>
<td>61.6040</td>
<td>7</td>
</tr>
</tbody>
</table>

TOTAL CASES = 67
## Appendix E

### Analysis of Variance and Orthogonal Polynomial Regression Statistics

**Pilot**

<table>
<thead>
<tr>
<th>MULTIPLE R</th>
<th>R SQUARE</th>
<th>ADJUSTED R SQUARE</th>
<th>STANDARD ERROR</th>
<th>ANALYSIS OF VARIANCE</th>
<th>REGRESSION</th>
<th>RESIDUAL</th>
<th>DF</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.37444</td>
<td>0.14116</td>
<td>0.12713</td>
<td>14.27676</td>
<td></td>
<td></td>
<td></td>
<td>65</td>
<td>13248.7773</td>
<td>203.52581</td>
<td>10.61285</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>STC ERROR</th>
<th>B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.23246</td>
<td>0.37444</td>
<td>0.06191</td>
<td>10.613</td>
</tr>
<tr>
<td>QUADR</td>
<td>0.212518</td>
<td>0.212518</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.15611</td>
<td>0.15611</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>QUARTIC</td>
<td>0.08344</td>
<td>0.08344</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.02574</td>
<td>0.02574</td>
<td>0.05919</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>STC ERROR</th>
<th>B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.23246</td>
<td>0.37444</td>
<td>0.06191</td>
<td>10.613</td>
</tr>
<tr>
<td>QUADR</td>
<td>0.212518</td>
<td>0.212518</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.15611</td>
<td>0.15611</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>QUARTIC</td>
<td>0.08344</td>
<td>0.08344</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.02574</td>
<td>0.02574</td>
<td>0.05919</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**FILE NAME** (CREATION DATE = 08/05/79)

**DEPENDENT VARIABLE...** PCT

**VARIABLE(S) ENTERED ON STEP NUMBER 1...** CLARITY

**MULTIPLE R**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>STC ERROR</th>
<th>B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.23246</td>
<td>0.37444</td>
<td>0.06191</td>
<td>10.613</td>
</tr>
<tr>
<td>QUADR</td>
<td>0.212518</td>
<td>0.212518</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.15611</td>
<td>0.15611</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>QUARTIC</td>
<td>0.08344</td>
<td>0.08344</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.02574</td>
<td>0.02574</td>
<td>0.05919</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**VARIABLE(S) ENTERED ON STEP NUMBER 2...** QUADR

**MULTIPLE R**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>STC ERROR</th>
<th>B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.23246</td>
<td>0.37444</td>
<td>0.06191</td>
<td>10.613</td>
</tr>
<tr>
<td>QUADR</td>
<td>0.212518</td>
<td>0.212518</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.15611</td>
<td>0.15611</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>QUARTIC</td>
<td>0.08344</td>
<td>0.08344</td>
<td>0.00395</td>
<td>0.000</td>
</tr>
<tr>
<td>CUBIC</td>
<td>0.02574</td>
<td>0.02574</td>
<td>0.05919</td>
<td>0.000</td>
</tr>
</tbody>
</table>

132
**Appendix E**  
(cont)

Pilot

---

**FILE NAME** (CREATION DATE = 08/03/79)

**MULTIPLE REGRESSION**  

**DEPENDENT VARIABLE:** PCT  

**VARIABLE(S) ENTERED ON STEP NUMBER 3:** QUARTIC

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>STANDARD ERROR</th>
<th>T-LEVEL</th>
<th>TOLERANCE-LEVEL</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>-0.497314</td>
<td>0.92379</td>
<td>1.06660</td>
<td>0.206</td>
<td></td>
</tr>
<tr>
<td>QUARTIC</td>
<td>0.113550</td>
<td>-0.03323</td>
<td>0.11929</td>
<td>0.328</td>
<td></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-0.77340</td>
<td>0.00030</td>
<td>0.00000</td>
<td>0.220</td>
<td></td>
</tr>
</tbody>
</table>

**ANALYSIS OF VARIANCE**  

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>STANDARD ERROR</th>
<th>T-LEVEL</th>
<th>TOLERANCE-LEVEL</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUBIC</td>
<td>-35.77754</td>
<td>-0.03933</td>
<td>0.00000</td>
<td>0.096</td>
<td></td>
</tr>
</tbody>
</table>

**VARIABLES NOT IN THE EQUATION**  

**VARIABLE** | **BETA** | **PARTIAL TOLERANCE** | **F**
|------------|----------|------------------------|---|

F-LEVEL OR TOLERANCE-LEVEL INSUFFICIENT FOR FURTHER COMPUTATION STASTICS WHICH CANNOT BE COMPUTED ARE PRINTED AS ALL NINES.
### Stepwise Regression Summary Table

**Pilot**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MULTIPLE R</th>
<th>R SQUARE</th>
<th>RSQ CHANGE</th>
<th>SIMPLE R</th>
<th>BETA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.37464</td>
<td>0.14036</td>
<td>0.00000</td>
<td>0.37464</td>
<td>-0.92379</td>
</tr>
<tr>
<td>QUAD</td>
<td>0.38520</td>
<td>0.14818</td>
<td>0.00802</td>
<td>0.38520</td>
<td>2.02323</td>
</tr>
<tr>
<td>QUARTIC</td>
<td>0.38903</td>
<td>0.15135</td>
<td>0.00297</td>
<td>0.38903</td>
<td>-0.77340</td>
</tr>
<tr>
<td>CONSTANT</td>
<td></td>
<td></td>
<td></td>
<td>0.36450</td>
<td></td>
</tr>
</tbody>
</table>

**FILE NAME** (CREATION DATE = 08/05/79)

**MULTIPLE REGRESSION**

**SUMMARY TABLE**
## Appendix G

### Description of Subpopulations

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>CPCE</th>
<th>VALUE LABEL</th>
<th>SUM</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>VARIANCE</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLN ENTIRE POPULATION</td>
<td>4491.0000</td>
<td>53.4643</td>
<td>12.6572</td>
<td>160.2035</td>
<td>6</td>
<td>841</td>
<td></td>
</tr>
<tr>
<td>CLA1RTY</td>
<td>91.</td>
<td>523.0000</td>
<td>57.3333</td>
<td>10.1266</td>
<td>102.7510</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>CLA1RTY</td>
<td>95.</td>
<td>557.0000</td>
<td>57.7500</td>
<td>16.5151</td>
<td>272.7523</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>CLA1RTY</td>
<td>99.</td>
<td>592.0000</td>
<td>60.2222</td>
<td>13.9226</td>
<td>192.4444</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

TOTAL CASES = 84
## Appendix H

### Analysis of Variance and Orthogonal Polynomial Regression Statistics

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>BETA</th>
<th>TOLERANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>-0.0287</td>
<td>0.0134</td>
</tr>
<tr>
<td>CUBIC</td>
<td>-0.1227</td>
<td>0.0001</td>
</tr>
<tr>
<td>QUANTIC</td>
<td>0.1478</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

F-LEVEL CR TOLERANCE - LEVEL INSUFFICIENT FOR FURTHER COMPUTATION STATISTICS WHICH CANNOT BE COMPUTED ARE PRINTED AS ALL NINES.
Appendix I

Stepwise Regression Summary Table

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MULTIPLE R</th>
<th>R-SQUARE</th>
<th>ADJUSTED R-SQUARE</th>
<th>SIMPLE R</th>
<th>BETA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLARITY</td>
<td>0.49531</td>
<td>0.24533</td>
<td>0.24533</td>
<td>0.49531</td>
<td>3.054612</td>
</tr>
<tr>
<td>QUAD</td>
<td>0.52590</td>
<td>0.25635</td>
<td>-0.01162</td>
<td>0.48761</td>
<td>-1.72718</td>
</tr>
<tr>
<td>CONSTANT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.026237</td>
</tr>
</tbody>
</table>
# Appendix J

95% Confidence Levels

## Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F Ratio</th>
<th>F Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>7</td>
<td>3951.593</td>
<td>278.730</td>
<td>3.736</td>
<td>0.0075</td>
</tr>
<tr>
<td>Within Groups</td>
<td>68</td>
<td>9493.814</td>
<td>139.619</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>13445.407</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Group Analysis

<table>
<thead>
<tr>
<th>GROUP</th>
<th>COUNT</th>
<th>PLAN</th>
<th>STANDARD DEVIATION</th>
<th>STANDARD DEVIATION</th>
<th>MINIMUM</th>
<th>MAXIMUM</th>
<th>95 PCT CONF INT FOR MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLP31</td>
<td>10</td>
<td>37.6092</td>
<td>11.1431</td>
<td>3.5470</td>
<td>22.3000</td>
<td>50.0000</td>
<td>28.4566 TO 45.1184</td>
</tr>
<tr>
<td>GLP32</td>
<td>7</td>
<td>53.4531</td>
<td>7.4342</td>
<td>1.6152</td>
<td>10.0000</td>
<td>61.0000</td>
<td>46.9357 TO 60.7495</td>
</tr>
<tr>
<td>GLP33</td>
<td>10</td>
<td>51.7732</td>
<td>10.3216</td>
<td>3.2633</td>
<td>33.0000</td>
<td>67.0000</td>
<td>43.5735 TO 58.4265</td>
</tr>
<tr>
<td>GLP53</td>
<td>6</td>
<td>55.0722</td>
<td>4.1787</td>
<td>1.3332</td>
<td>50.0000</td>
<td>80.0000</td>
<td>58.0118 TO 64.4327</td>
</tr>
<tr>
<td>GLP76</td>
<td>12</td>
<td>57.0900</td>
<td>16.5411</td>
<td>4.7575</td>
<td>27.0000</td>
<td>78.0000</td>
<td>66.7968 TO 77.7412</td>
</tr>
<tr>
<td>GLP77</td>
<td>10</td>
<td>54.6072</td>
<td>13.3452</td>
<td>4.1753</td>
<td>36.0000</td>
<td>77.0000</td>
<td>66.6660 TO 72.4317</td>
</tr>
<tr>
<td>TOTAL</td>
<td>76</td>
<td>53.5000</td>
<td>13.2388</td>
<td>4.1753</td>
<td>36.0000</td>
<td>77.0000</td>
<td>44.5549 TO 70.1455</td>
</tr>
</tbody>
</table>

### Random Effects Model

| RANDOM EFFECTS MODEL | 2.6410 | 67.2551 | 59.7049 |

## Tests for Homogeneity of Variances

- Cochran C = Max. Variance/Sum of Variances = 0.2634, P = 0.142 (APPROX.)
- Hartley F = Max. F = 4.1753, P = 0.032
- Maximum Variance / Minimum Variance = 10.032

---

138
<table>
<thead>
<tr>
<th>Observation</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.3</td>
<td>4.5</td>
<td>6.7</td>
</tr>
<tr>
<td>2</td>
<td>3.4</td>
<td>5.6</td>
<td>7.8</td>
</tr>
<tr>
<td>3</td>
<td>4.5</td>
<td>6.7</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Appendix K

Scheme for the .05 level
Appendix L

Scheffe for the .100 level

<table>
<thead>
<tr>
<th>FILE NAME</th>
<th>CREATION DATE</th>
<th>ONE WAY</th>
</tr>
</thead>
</table>

**MULTIPLE RANGE TEST**

Scheffe (Holm)

RANGES FOR THE .100 LEVEL

5.03 5.03 5.03 5.03 5.03 5.03

The ranges above are table ranged; the value actually compared with H(K(1) + | K(1))

([0]) denotes pairs of groups significantly different at the .0100 level

<table>
<thead>
<tr>
<th>MEAN</th>
<th>GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.000</td>
<td>G 1</td>
</tr>
<tr>
<td>31.071</td>
<td>G 2</td>
</tr>
<tr>
<td>35.222</td>
<td>G 3</td>
</tr>
<tr>
<td>59.130</td>
<td>G 4</td>
</tr>
<tr>
<td>57.507</td>
<td>G 5</td>
</tr>
<tr>
<td>59.131</td>
<td>G 6</td>
</tr>
<tr>
<td>57.222</td>
<td>G 7</td>
</tr>
</tbody>
</table>

140
Appendix M

Reading Passage: Control 1--Scrambled Text

Kentucky born Hardin 12 I 1809 County was February in. Virginia parents families second undistinguished I families perhaps should my in of were say both were. Tenth mother year my my in who a from Hanks was died whom of now some County reside Adams Illinois others County im Macon named family others. Grandfather 1781 Virginia paternal Abraham to or 17182 Lincoln about Kentucky by where killed the or a later was Indians he year not but in when battle in to forest the open was he he two laboring was from farm a.

Death at the my father father of six his old without was grew and years he literally up education. Kentucky he Indiana from moved when to seven was I. Home reached about we new our state time the the Union into the came. Region was bears with wild was it a many and animals still woods wild in other the and. There up grew I. So-called were some no but there ever qualification schools was teacher a required reading of adding beyond and writing other. Stranger Latin a reside if time in for neighborhood understand happened was wizard looked as on a a the in he to. Absolutely excite there education was for ambition nothing to. Course age when did much not I of came I much know did. I but still all somehow read that could add was write and. Advance store little this made I up under pick-ed necessity have now of pressure education I have upon the of.

Raised was work which farm I I until two-twenty was con-tinued to I. Illinois cams one-twenty at to I. New Salem remained in I Illinois store year clerk for as a for a. Hawk came War then the Black captain was I success of a more gave which pleasure me than since had any have I volunteers of. Campaign into ran legislature for the I defeated year and was the cans (1832) rejected time people was only I have been ever the by. Three next in the and the biennial was legislature elected I elections succeeding to the. Candidate afterward I a not was period I law that legislative studied Springfield and to moved it practice during to. 1846 in to the elected I lower was Congress of house. Candidate a for was re-election not I. 1854 1849 for to law assiduously more before than practiced ever I. Losing politics was when interest repeal the I Compromise the of again me in Missouri aroused. Known done well what since I pretty is then have.

Description of if thought personal desirable me any is said be am that feet inches height in six four I may am it nearly; flesh weighing and in one lean average eighty hundred
Control 1--Continued

on the pounds. Marks no brands have other or I.
Appendix N

Reading Passage: Treatment I (1:2)

I was xif February 12, yuc, in Hardin Xow, Kentucky. My zec were both kaq in Virginia, qit undistinguished families—xug families perhaps Vjo should say. Diz mother, who qu in my qak year, was ciy a family xir Hanks, some qos whom now quc in Adams Yav, and others jez Macon County, Qit. My paternal wic, Abraham Lincoln, wiq, about 1781 puj 1782, from Xew to Kentucky, where zuj was killed qon the Indians fuq year or qim later, not yev battle, but yib he was pio to open sdr farm in wuk forest.

My jez, at the yaz of his qel, was six wqu old, and zuc grew up qow without education. Siw moved from Zaj to Indiana yiw I was qiw. We reached ceh new home xig time coz state came vef the Union. Xib was a jih region, with yig bears and jiy animals still yoq the woods. Ziy grew up xug. There were vjig so-called schools, jow no other qor was ever kuq of a hib beyone "reading, xol, and adding." Zqa a stranger zic to understand Huw happened to gaq foq a kif in the qoz, he was pliv on as my wizard. There cuh absolutely nothing tuy excite ambition xoj education. Of wera3 when I xil of age, Sic did not veg much. Still riw, I could quXj, write, and qiw, but that qaw all. The xes advance I cuv now made zuw this store quh education, I qw picked up woq the pressure vaq necessity.

I cuj raised to to piw work, which xas continued until Kuv was twenty-two. Xau twenty-one, I cil to Illinois. Hec in Miy Hil, Illinois for vuq year as yuv clerk in zin store. Then lew Black Hawk Xiq came; I xef made a qaw of volunteers, xaf success which nav me more qex than any Vez have had dof. I went xaq the campaign, zac for the mej the same qoh (1832), and juc defeated – the xet time I gik ever been xit by the vud. In the xad and the biw succeeding biennial wev I was qut to the piy. I was tej a candidate muy. During that kic period, I coj law, and hiw to Springfield nux practice it. Jiw 1846 I vus elected to xeq lower house forq Congress. I xiz not a x1w for re0election. Qar 1849 to yoz I practiced qej more assiduously qat ever before. Zob was losing leq in politics yuj the repeal biw the Missouri Wij aroused me xok. What I xiq done since xad is pretty xes known.

If zew personal description kux me is heq desirable, it xal be said giz I am jiz six feet, jud inches in xul; lean in caj, weighing one xey and eighty jok on the vay; I have xaw dark complexion, veg coarse black vog and gray jiq. I have xok other marks hos brands.
Appendix O

Reading Passage: Treatment 2 (1:4)

I was born February xif, 1809 in Hardin County, Xow. My parents were both zuc in Virginia, of undistinguished kaz - second families, perhaps I quz say. My mother, who xug in my tenth year, vjo from a family named XaI, some of whom now diz in Adams County, and qel in Macon County, Illinois. Qak paternal grandfather, Abraham Lincoln, cyz about 1781 or 1782, xir Virginia to Kentucky, where qos was killed by the Qac a year or two qos, not in battle, but yay he was laboring to zez a farm in the git.

My father, at the wic of his father, was wug years old, and he xew up laterally without education. Zug moved from Kentucky to Cim when I was juz. We reached our new yev about the time the yib came into the Union. Qim was a wild region, feq many bears and other qim animals still in the puj. I grew up there. Sdz were some so-called schools, wuk no other qualification was jez of a teacher beyond "yaz, writing, and adding." If qel stranger supposed to understand Wqu happened to reside for qow time in the neighborhood, siw was looked on as jeq wizard. There was absolutely yiw to excite ambition for xzo. Of course, when I caz of age, I did vef know much. Still, somehow, Ziy could read, write, and zqa but that was all. Jih little advance I have kuj now upon this store hib education, I have picked jow under the pressure of xuk.

I was raised to vjq work, which I continued kuj I was twenty-two. At qor-qar, I came to Illinois. Zic remained in New Salem, Illinois kuj a year as a huv in a store. Then miy Black Hawk War came; Qoz was elected a captain piv volunteers; a success which gaq me more pleasure than qix I have had since. Yiq went into the campaign, wem for the legislature the ziw year (1832), and was cuh - the only time I xoI ever been re- jected by yuw people. In the next qow the three succeeding biennial xes I was ejected to tgz legislature. I was not a qaw afterward. During that legislative lum, I studied law, and riw to Springfield to practice xet. In 1846 I was lwu to the lower house qox Congress. I was not cIl candidate for re-election. From juc to 1854 I practiced xiq more assiduously than ever taj. I was losing interest nav politics when the repeal zin the Missouri Compromise aroused qar again. What I have yoz since then is pretty xag known.

If any personal xaf of me is thought jiw, it may be said dof I am nearly six vud in height; lean in xuv, weighing one
hundred and kic pounds on the average; Piy have a dark complexion, qej coarse black hair and caj eyes. I have no niq marks or brands.
Appendix P

Reading Passage: Treatment 3 (1:6)

I was xif February 12, 1809, in Hardin County, Xow. My parents were born in Kaq, of undistinguished families - second families, perhaps Voj should say. My mother, who died xir my tenth year, was from a qit named Hanks, some of whom now xic in Adams County, and others in Wic County, Illinois. My paternal grandfather, Abraham Wuq, emigrated, about 1781 or 1782, from Kir to Kentucky, where he was killed puj the Indians a year or two qim, not in battle, but when he fuq laboring to open a farm in new forest.

My father, at the death jez his father, was six years old, sdr he grew up literally without education. Qel moved from Kentucky to Indiana when Zaj was seven. We reached our new yib about the time the state came yaz the Union. It was a wild qow, with many bears and other wild yiq still in the woods. I grew xiq there. There were some so-called schools, coz no other qualification was ever required jin a teacher beyond "reading, writing, and diz." If a stranger supposed to understand Yev happened to reside for a time cjm the neighborhood, he was looked on xib a wizard. There was absolutely nothing fuq ex­cite ambition for education. Of course, siw' I came of age, I did xoj know much. Still, somehow, I could viq, write, and add, but that was piv. The little advance I have now zic upon this store of education, I xes picked ip under the pressure of qew.

I was raised to farm work, xic I continued until I was twenty-two. Saq twenty-one, I came to Illinois. I kuj in New Salem, Illinois for a year veq a clerk in a store. Then xoj Black Hawk War came; I was voq a captain of volunteers, a success luv gave me more pleasure than any Woq have had since. I went into piy campaign, ran for the legislature the zin year (1832), and was defeated - the guh time I have ever been re­jected gur the people. In the next and juc three succeeding biennial elections I was zob to the legislature. I was not xaf candidate afterward. During that legislative period, Hec studied law, and moved to Springfield xiq practice it. In 1846 I was xad to the lower house of Congress. Xok was not a candidate for re-election. Gik 1849 to 1854 I practiced law quit assiduously than ever before. I was kic interest in poli­tics when the repeal xaf the Missouri Compromise aroused me again. Xiw I have done since then is youz well known.

If any personal description xet me is thought desirable, it may biw said that I am nearly six vud in height; lean in
flesh, weighing vay hundred and eithty pounds on the mif; I have a dark complexion, with wijn black hair and gray eyes. I jif no other marks or brands.
Appendix Q

Reading Passage: Treatment 4 (1;8)

I was born February yuc, 1809, in Hardin County, Kentucky. My parents were now born in Virginia, of undistinguished families - second families, zuc I should say. My mother, who died in my tenth year, was from a family named Hanks, xif of whom now reside in Adams County, and kaq in Macon County, Illinois. My paternal grandfather, Abraham Voj, emigrated, about 1781 or 1782, from Virginia to Qim, where he was killed by the Indians a wid or two later, not in battle, but when diz was laboring to open a farm in the new.

My father, at the death of his father qel six years old, and he grew up literally quc education. He moved from Kentucky to Indiana when Qos was seven. We reached our new home about wug time the state came into the Union. It fuq a wild region, with many bears and other xew animals still in the woods. I grew up tej. There were some so-called schools, but no other yev was ever required of a teacher beyond "reading, qim, and adding." If a stranger supposed to understand Cij happened to reside for a time in the sdr, he was looked on as a wizard. There wuk absolutely nothing to excite ambition for education. Of qel, when I came of age, I did not jex much. Still, somehow, I could read, write, and add, yaz that was all. The little advance I have wuq made upon this store of education, I have pif up under the pressure of necessity.

I was zuc to farm work, which I continued until I qow twenty-two. At twenty-one, I came to Illinois. I siw in New Salem, Illinois for a year as a zaj in a store. Then the Black Hawk War yiw; I was elected a captain of volunteers, a qiw which gave me more pleasure than any I ceh had since. I went into the campaign, ran feq the legislature the same year (1832), and was wiq - the only time I have ever been rejected vef the people. In the next and three jiy biennial elections I was elected to the legislature. Yoq was not a candidate afterward. During the legislative vjg, I studied law, and moved to Spring- field to xol it. In 1846 I was elected to the guq house of Congress. I was not a candidate zqa re-election. From 1849 to 1854 I practiced law zes assiduously than ever before. I was losing interest piv politics when the repeal of the Missouri Compromise qoz me again. What I have done since then yib pretty well known.

If any personal description of joh is thought desirable, it may be said that Xiv am nearly six feet, four inches in height; zes in flesh, weighing one hundred and eighty pounds
Appendix Q--Continued

I'm the average; I have a dark complexion, with xad black hair and gray eyes. I have no jig marks or brands.
Appendix R

Reading Passage: Treatment 5 (1:10)

I was born February 12, 1809, in Harding County, Kentucky. My parents were both born in Virginia, of undistinguished families — second now, perhaps I should say. My mother, who died in my ninth year, was from a family named Hanks, some of whom now reside in Adams County and others in Macon County, Illinois. My paternal grandfather, Abraham Lincoln, emigrated, about 1781 or 1782, from Kentucky to Kentucky, where he was killed by the Indians about two later, not in battle but when he was trying to open a farm in the forest.

My father, at his death of his father, was six years old, and he grew up literally without education. He moved from Kentucky to Indiana when I was seven. We reached our new home about the year the state came into the Union. It was a wild camp, with many bears and other wild animals still in the region. I grew up there. There were some so-called schools, but no other qualification was ever required of a teacher beyond "reading, cipher, and adding." If a stranger supposed to understand Latin happened to reside for a time in the neighborhood, he was looked on as a wizard. There was absolutely nothing to excite ambition in education. Of course, when I came of age, I did not know much. Still, somehow, I could read, write, and add, which was all. The little advance I have now made is from this store of education. I have picked up under the xes of necessity.

I was raised to farm work, which I did until I was twenty-two. At twenty-one, I came to Illinois. I remained in New Salem, Illinois for a year as a clerk in a store. Then the Black Hawk War came; I was a captain of volunteers, a success which gave me more respect than any I have had since. I went into the doff, ran for the legislature the same year (1832), and was qew — the only time I have ever been rejected by the xil. In the next and the three succeeding biennial elections I was qew to the legislature. I was not a candidate afterward. During the legislative period, I studied law, and moved to Springfield to xaf it. In 1846 I was elected to the lower house xox Congress. I was not a candidate for re-election. From 1849 xes 1854 I practiced law more assiduously than ever before. I vud losing interest in politics when the repeal of the Missouri xin aroused me again. What I have done since then is xun well known.

If any personal description of me is desired, chip may be said that I am nearly six feet, four inches in height;
lean in flesh, weighing one hundred and eighty hac on the average; I have a dark complexion, with coarse xad hair and gray eyes. I have no other marks or xoj.
Appendix S

Reading Passage: Treatment 6 (1:12)

I was xif February 12, 1809, in Hardin County, Kentucky. My parents were both born now Virginia, of undistinguished families — second families, perhaps I should say. My mother, yic died in my tenth year, was from a family named Hanks, some zu which now reside in Adams County, and others in Macon County, Illinois. Kaj paternal grandfather, Abraham Lincoln, emigrated, about 1781 or 1782, from Virginia to Oit, where he was killed by the Indians a year or two later, xew in battle, but when he was laboring to open a farm in qim forest.

My father, at the death of his father, was six years yaz, and he grew up letterally without education. He moved from Kentucky to Xir when I was seven. We reached our new home about the time wic state came into the Union. It was a wild region, with many yev and other wild animals still in the woods. I grew up there. Sdr were some so-called schools, but no other qualification was ever required of wuk teacher beyond "reading, writing, and adding." If a stranger supposed to understand Jez happened to reside for a time in the neighborhood, he was looked yig as a wizard. There was absolutely nothing to excite ambition for education. Vjq course, when I came of age, I did not know much. Qow, somehow, I could read, write, and add, but that was all. The yaz advance I have now made upon this store of education, I have zag up under the pressure of necessity.

I was raised to farm work, was I continued until I was twenty-two. At twenty-one, I came to Illinois. Xig remained in New Salem, Illinois for a year as a clerk in a xib. Then the Black Hawk War came; I was elected a captain of gwe, a success which gave me more pleasure than any I have had jiw. I went into the campaign, ran for the legislature the same year (biw), and was defeated - the only time I have ever been rejected by xic people. In the next and the three succeeding biennial elections I was zag to the legislature. I was not a candidate afterward. During that legislative xaf, I studied law, and moved to Springfield to practice it. In 1846 Xet was elected to the lower house of Congress. I was not a yuv for re-election. From 1849 to 1854 I practiced law more assiduously than coz before. I was losing interest in politics when the repeal of the Xad Compromise aroused me again. What I have done since then is pretty fog known.

If any personal description of me is thought desirable, it may vus said that I am nearly six feet, four inches in
Appendix S--Continued

height; lean but flesh, weighing one hundred and eighty pounds on the average; I have dark complexion, with coarse black hair and gray eyes. I have no leg marks or brands.
A p p e n d i x  T

Reading Passage: Treatment 7 (1:14)

I was born February 12, 1809, in Hardin County, Kentucky. My parents were both born in Xif, of undistinguished families - second families, perhaps I should say. My mother, who died in yuc year was from a family named Hanks, some of whom now reside in Zuc County, and others in Macon County, Illinois. My paternal grandfather, Abraham Lincoln, emigrated, about kaq or 1782, from Virginia to Kentucky, where he was killed by the Indians a year qit two later, not in battle, but when he was laboring to open a farm xuq the forest.

My father, at the death of his father, was six years old, vjo he grew up literally without education. He moved from Kentucky to Indiana when I gel seven. We reached our new home about the time the state came into the Xir. It was a wild region, with many bears and other wild animals still in qak woods. I grew up there. There were some so-called schools, but no other zuj was ever required of a teacher beyond "reading, writing, and adding." If a stranger wuk to understand Latin happened to reside for a time in the neighborhood, he was qim on as a wizard. There was absolutely nothing to excite ambition for education. Of yev, when I came of age, I did not know much. Still, somehow, I could yaz, write, and add, but that was all.

The little advance I have now made sdr this store of education, I have picked up under the pressure of necessity.

I qow raised to farm work, which I continued until I was twenty-two. At twenty-one, I zaq to Illinois. I remained in New Salem, Illinois for a year as a clerk in kuj store. Then the Black Hawk War came; I was elected a captain of volunteers, xuq success which gave me more pleasure than any I have had since. I went yib the campaign, ran for the legislature the same year (1832), and was defeated - the ceh time I have ever been rejected by the people. In the next and the meses succeeding biennial elections I was elected to the legislature. I was not a candidate voq. During the legislative period, I studied law, and moved to Springfield to practice it. Piw 1846 I was elected to the lower house of Congress. I was not a zic for re-election. From 1849 to 1854 I practiced law more assiduously than ever before. Xiq was losing interest in politics when the repeal of the Missouri Compromise aroused me xed. What I have done since then is pretty well known.

If any personal description jih me is thought desirable, it may be said that I am nearly six feet, kic inches in height; lean in flesh, weighing one hundred and eighty pounds on the
Appendix T—Continued

xoj; I have a dark complexion, with coarse black hair and gray eyes. I have wij other marks or brands.
Appendix U

Reading Passage: Control 2 (Unaltered Text)

I was born February 12, 1809, in Hardin County, Kentucky. My parents were both born in Virginia, of undistinguished families — second families, perhaps I should say. My mother, who died in my tenth year, was from a family named Hanks, some of whom now reside in Adams County, and others in Macon County, Illinois. My paternal grandfather, Abraham Lincoln, emigrated, about 1781 or 1782, from Virginia to Kentucky, where he was killed by the Indians a year or two later, not in battle, but when he was laboring to open a farm in the forest.

My father, at the death of his father, was six years old, and he grew up literally without education. He moved from Kentucky to Indiana when I was seven. We reached our new home about the time the state came into the Union. It was a wild region, with many bears and other wild animals still in the woods. I grew up there. There were some so-called schools, but no other qualification was ever required of a teacher beyond "reading, writing, and adding." If a stranger supposed to understand Latin happened to reside for a time in the neighborhood, he was looked on as a wizard. There was absolutely nothing to excite ambition for education. Of course, when I came of age, I did not know much. Still, somehow, I could read, write, and add, but that was all. The little advance I have now made upon this store of education, I have picked up under the pressure of necessity.

I was raised to farm work, which I continued until I was twenty-two. At twenty-one, I came to Illinois. I remained in New Salem, Illinois for a year as a clerk in a store. Then the Black Hawk War came; I was elected a captain of volunteers, a success which gave me more pleasure than any I have had since. I went into the campaign, ran for the legislature the same year (1832), and was defeated — the only time I have ever been rejected by the people. In the next and the three succeeding biennial elections I was elected to the legislature. I was not a candidate afterward. During that legislative period, I studied law, and moved to Springfield to practice it. In 1846 I was elected to the lower house of Congress. I was not candidate for re-election. From 1849 to 1854 I practiced law more assiduously than ever before. I was losing interest in politics when the repeal of the Missouri Compromise aroused me again. What I have done since then is pretty well known.

If any personal description of me is thought desirable, it may be said that I am nearly six feet, four inches in height,
lean in flesh, weighing one hundred and eighty pounds on the average; I have a dark complexion, with coarse black hair and gray eyes. I have no other marks or brands.
Appendix V

Cloze Test of Comprehension

I was ______ February 12, 1809, in ______ County, Kentucky. My parents ______ both born in Virginia, ______ undistinguished families - second families ______ I should say. My ______ who died in my ______ year, was from a ______ named Hanks, some of ______ now reside in Adams ______ and others in Macon ______ Illinois. My paternal grandfather, _______ Lincoln, emigrated, about 1781 ______ 1782, from Virginia to _______ where he was killed ______ the Indians a year ______ two later, not in _______ but when he was ______ to open a farm _______ the forest.

My father, _______ the death of his _______ was six years old, _______ he grew up literally _______ education. He moved from _______ to Indiana when I _______ seven. We reached our _______ home about the time _______ state came into the _______. It was a wild _______ with many bears and _______ wild animals still in _______ woods. I grew up _______. There were some so-called _______ but no other qualification _______ ever required of a _______ beyond "reading, writing, and _______." If a stranger supposed _______ understand Latin happened _______ reside for a time _______ the neighborhood, he was _______ on as a wizard.
Appendix V--Continued

was absolutely nothing to ambition for education. Of when I came of, I did not know much. Still, somehow, I could write, and add, but was all. The little I have now made this store of education, have picked up under pressure of necessity.

I raised to farm work, I continued until I twenty-two. At twenty-one, I to Illinois. I remained New Salem, Illinois for a as a clerk in store. Then the Black War came; I was a captain of volunteers, success which gave me pleasure than any I had since. I went the campaign, ran for legislature the same year, and was defeated - the time I have ever rejected by the people. the next and the succeeding biennial elections I elected to the legislature. was not a candidate.

During that legislative period, studied law, and moved Springfield to practice it. 1846 I was elected the lower house of. I was not a for re-election. From 1849 1854 I practiced law assiduously than ever before.
Appendix V--Continued

losing interest in ______ when the repeal of ______ Missouri Compromise aroused me _______. What I have done ______ then is pretty well _______.

If any personal description ______ me is thought desirable, ______ may be said that ______ am nearly six feet, ______ inches in height; lean ______ flesh, weighing one hundred ______ eighty pounds on the ______; I have a dark ______, with coarse black hair ______ gray eyes. I have ______ other marks or brands.