Techniques for Increasing Writing and Reading Proficiency of International Students

Abatah D. Daher

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses

Part of the Educational Psychology Commons, and the Language and Literacy Education Commons

Recommended Citation
https://scholarworks.wmich.edu/masters_theses/1106
TECHNIQUES FOR INCREASING WRITING
AND READING PROFICIENCY OF
INTERNATIONAL STUDENTS

by

Abatah D. Daher

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
December 1989
TECHNIQUES FOR INCREASING WRITING AND READING PROFICIENCY OF INTERNATIONAL STUDENTS

Abatah D. Daher, M.A.
Western Michigan University, 1989

The purpose of this study was to examine and develop a program to increase the language proficiency of international students.

The subjects for this study consisted of 44 international students enrolled in the course entitled, "Effective College Reading for International Students." The subjects were divided into two groups, the experimental group and the control group.

Gates and MacGinitie Reading Survey E (1978) was used in assessing language proficiency. Speed and accuracy, vocabulary, and comprehension scores were obtained for the experimental and control groups. An analysis using the Fry formula (1977) to identify the grade level score for their stories was performed each week and the grade level score provided.

The analysis of the data using the t-test revealed significant gains. The results of writing performance also showed significant differences in international students between the beginning and the end of the program.
ACKNOWLEDGEMENTS

This paper is dedicated to my family without whom it never would have been completed. My parents provided much support and encouragement throughout my graduate career. To my son, Tarek, and my daughter, Dena, I offer my love and thanks for this among the other changes they have brought into my life. But the greatest love and thanks goes to my wife and friend, Faten, who loved and supported me through the difficult times, and who also had to endure my long hours away from home and long hours of studying. Faten suffered through all of this and still showed me much love and support.

Special thanks goes to my committee chair, Dr. Dale Brethower, for his cooperation, encouragement, and guidance throughout this study. I also thank the members of my advisory committee, Dr. Richard Malott and Dr. Neil Kent.

Appreciation is also extended to Dr. Bruce Lloyed for his aid in helping to obtain the subjects for this investigation. Additional acknowledgement is extended to Mrs. Suzanne Sippola and Ms. Sue Kent for the typing of the form of this manuscript.

Abatah D. Daher
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book. These are also available as one exposure on a standard 35mm slide or as a 17" x 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700   800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Techniques for increasing writing and reading proficiency of international students

Daher, Abatah Dabaan, M.A.
Western Michigan University, 1989
TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................ ii
LIST OF TABLES.............................................. v
LIST OF FIGURES........................................... vii
CHAPTER
   I. INTRODUCTION........................................ 1
      Statement of the Problem............................. 17
      Research Questions.................................. 17
   II. DESIGN AND METHOD.................................. 19
      Research............................................. 19
      Setting.............................................. 19
      Subjects............................................. 19
      Apparatus and Materials............................. 20
      Methodology......................................... 20
      Experimental Design................................ 22
   III. RESULTS............................................ 24
      Testing the Hypotheses............................... 25
      Hypothesis One...................................... 25
      Hypothesis Two...................................... 32
      Findings of Secondary Analysis Data............... 41
LIST OF TABLES

1. A Comparison of the Mean Scores for the Experimental Group in Speed and Accuracy for the Pretest and Posttest. .......................... 26

2. A Comparison of the Mean Scores for the Control Group in Speed and Accuracy for the Pretest and Posttest. .......................... 27

3. A Comparison of the Mean Scores for the Experimental Group for the Vocabulary Pretest and Posttest. ....................... 28

4. A Comparison of the Mean Scores for the Control Group in Vocabulary for the Pretest and Posttest. .......................... 29

5. A Comparison of the Mean Scores for the Experimental Group on Comprehension for the Pretest and Posttest. ....................... 30

6. A Comparison of the Mean Scores for the Control Group on Comprehension for the Pretest and Posttest. ....................... 31

7. A Comparison of the Mean Scores for the Experimental and Control Groups for Vocabulary on the Pretest ....................... 36

8. A Comparison of the Mean Scores for the Experimental and Control Groups for Vocabulary on the Posttest ....................... 37

9. A Comparison of the Mean Scores for the Experimental and Control Groups for Speed and Accuracy on the Pretest ....................... 38

10. A Comparison of the Mean Scores for the Experimental and Control Groups for Speed and Accuracy on the Posttest ....................... 39

11. A Comparison of the Mean Scores for the Experimental and Control Groups for Comprehension on the Pretest ....................... 40
List of Tables—Continued

12. A Comparison of the Mean Scores for the Experimental and Control Groups for Comprehension on the Posttest .................. 41

13. The Percentage of Posttest Minus Pretest Divided by Pretest for Experimental Group ........................................ 45

The Percentage of Posttest Minus Pretest Divided by Pretest for Control Group ........................................ 45

14. The Percentage of Items Correct Divided by Total Items for Experimental Group .................. 48

15. The Percentage of Items Correct Divided by Total Items for Control Group .................. 51
LIST OF FIGURES

1. Data From Experimental Group ....................... 46
2. Data From Control Group ........................... 47
3. Data From Experimental Group Pretest ............... 49
4. Data From Experimental Group Posttest ............. 50
5. Data From Control Group Pretest .................... 52
6. Data From Control Group Posttest .................. 53
CHAPTER I

INTRODUCTION

Language skills are usually defined in terms of the receptive language activities of reading and listening and the expressive language activities of speaking and writing. Language seems to be the designation which best describes the activities of speaking, writing, listening, and reading and best labels an instructional program which focuses developing abilities in the language activities (Greene & Petty, 1971).

Language skills are mandatory for international students. Language is the basic tool of communication and the specific skills, abilities and attitudes needed by students for effective communication must be learned directly and thoroughly. The student acquires a mastery of language by using it. The university curriculum must be rich in content and activities which stimulate and promote extensive language use. Therefore, a language program must do as much as possible to assure practice with a variety of language skills.

As an international student, I am aware of the problems of learning a new language. Further, my experience was not a unique one. After extensive questioning of other foreign students, I discovered that they all experienced a language
difficulty to a greater or lesser degree.

This study was designed to increase reading, writing and other language proficiency such as speed and accuracy, as well as vocabulary and comprehension.

Martin and Gail (1979) researched the development of English proficiency of 83 foreign students who were living in English-speaking American homes during their 14-week intensive formal language program. Statistical comparison with nonhomestay students revealed that homestay students achieved significantly higher course grades and higher scores on the test of English as a foreign language. However, significant differences in the test of English as a foreign language overall and reading comprehension scores seemed to be more a function of native language than either home or nonhome-residency. Home-residency seemed to be the factor most correlated with grades. The study lent support to the assumption that the best way to acquire language competency is to live in an environment where the language of study is used. It is speculated that the success of homestay students may be the result of practicing the second language; also, it is possible the student who chooses a homestay is likely to be outgoing and highly motivated to learn a foreign language.

Alderson and Urguhart (1984) studied the effects of foreign students' knowledge of a particular academic area on their comprehension of written texts in English. They found
consistent differences in comprehension between groups and among texts. Some slight discrepancies with expected results were found, but overall the studies confirmed the importance of prior knowledge in reading comprehension in English.

Harvey (1981) argued that standard assumptions and teaching strategies must be discarded when working with international students on the university level. The problems faced by international students are exacerbated by archaic curriculum designs for reading courses. Students have often been expected to complete major literary works after only three semesters in the language. They were expected to progress from a language novice unused to intensive work and reading to a mature reader able to handle a vocabulary of 80,000 to 120,000 words. In order to curb attrition, a new curriculum was designed which attended to foreign students' needs. Assessment was made of student cognitive styles and a personal system of instruction for listening and reading comprehension was designed. The latest strategies were taught to the foreign students in the course. No ability to read in the target language was presupposed. Rate-alteration technology was applied to the reading-listening curriculum in the textual phase of a modified directed reading approach. The result of this combination of methods was better reading during a one-semester course than listening comprehension.
Recent reviews of vocabulary training research (Jenkins & Penny, 1981; Mezynski, 1983) indicated that many studies succeed in increasing word knowledge. Several studies resulted in improvement in reading and writing that could be attributed to vocabulary instruction (Beck, Perfetti, & McKeown, 1982).

Positive results were also observed when remedial reading programs were used (Schwartz, 1977) and when college students were provided monetary rewards for increased study time (Bristol & Sloane, 1974).

Wixson (1986) randomly assigned students from two schools to one of eight groups according to word level (central or noncentral), method of instruction (dictionary or concept), and story (king or cave). After the data were collected, there were exactly 15 subjects in each group. After all of the adjustments had been made to each story, including the vocabulary, each story was approximately 1,500 words. The result of this was that the cave story had 59 central and 88 noncentral words; and the king story had 44 central and 89 noncentral words.

The study determined the effects of preteaching vocabulary of differing levels of importance to a text using two different methods of instruction on students' comprehension of basal stories. One hundred and twenty average and above-average readers participated in two group sessions. During the first session, small groups of subjects
received a vocabulary list of one set of five words (central or noncentral), using one of two methods of instruction (dictionary or concept) for one of two stories. Twenty-four hours later the second session was held, and two comprehension measures (recall and questions) and two vocabulary measures (definition and example) were administered. The results of the word-level treatment indicated that preteaching unfamiliar vocabulary enhanced students' comprehension of story ideas that were related to the instructed vocabulary regardless of level of importance.

In another study, central/noncentral vocabulary words were selected from unfamiliar words in each story and identified by using the index of vocabulary familiarity (Dale & O'Rourke, 1976). This index reports the percentage of students by grade level who knew specific meanings of the entry words in large-scale testing. The unfamiliar words were underlined, and 45 adults in graduate-level education courses were asked to read the stories and indicate which of the underlined words were important ideas in the story and which were less central or important. The five words that the adult readers agreed were most central and the five that they agreed were least central were selected for use in this experiment. Agreement for each of the selected words was 70% or greater. Furthermore, it was determined that all of the central words were located in central idea units in the stories, and that the noncentral words were located only in
noncentral idea units.

The level of the instructed vocabulary had a reliable effect on both word knowledge and text comprehension. Evidence for this comes from the presence of significant word level by response level interaction effects for both of the vocabulary measures and the comprehension question measure, and the absence of any significant effects for interactions with either method or story. Specifically, students who received instruction on central words learned more central and noncentral vocabulary words. Furthermore, students who received instruction on noncentral words learned more noncentral vocabulary and understood more story ideas related to noncentral vocabulary than students who received instruction on central words. These results suggest that preteaching unfamiliar vocabulary training enhances students' comprehension of story ideas related to the instructed vocabulary regardless of level of importance. The results indicate that comprehension questions designed to evaluate students' understanding of story ideas that are more sensitive measure of students' comprehension than a general recall measure.

Follow-up tests to the Word Level and Response Level interaction indicated that the subjects who received instruction on the central words performed in a similar manner on both central and noncentral example items (p > .05), whereas, the subjects who received instruction on the
noncentral words performed better on the noncentral than the central example items. Follow-up tests also indicated that the subjects who were taught central vocabulary words answered correctly more central vocabulary words answered correctly more central example items than the subjects who were taught the noncentral words, whereas the subjects who were taught the noncentral words answered more noncentral example items correctly than those who were taught central words (Wixson, 1986).

Two studies have been carried out on the measurement of general reading comprehension. Jenkinson (1957) correlated cloze scores with the results of the "vocabulary" and "level of comprehension", and obtained correlations of .78 and .73, respectively. Rankin (1957), using the section as a criterion of general reading skill, obtained the following correlations (uncorrected for attenuation) with cloze scores: story comprehension .29, vocabulary .68, and paragraph comprehension .60.

According to Barnas, Ginther, and Cochran (1989) good and poor comprehenders use knowledge and purpose in learning vocabulary from context. It was expected that good readers would do better at learning new words from reading context than poor readers. Subjects who scored at or above 60% correct were considered good reading comprehenders. Those who scored 40% or below were considered poor comprehenders.
Durkin (1979) found that in 4,469 minutes of reading instruction the teacher spent only 19 minutes teaching vocabulary. Nor were teachers providing much vocabulary instruction during content lessons, such as social studies. As a rule, basal reading programs do not emphasize vocabulary instruction. One major program provides no formal vocabulary instruction and another target only 300 "new" words for instruction, many of which average students already know. In reviewing the vocabulary teaching procedures of basal reading programs, Durkin found fault even with the "best case" instruction. Under these circumstances students encountered new vocabulary only three times: (1) before reading a story when they were told meanings for unfamiliar words, (2) when the words appeared in the story, and (3) in an independent seatwork exercise.

Duffy, et al. (1986) pointed out that reading comprehension emphasizes the strategic role of the reader. The goal of this study was to determine whether, given skills prescribed in a mandated basal reading series, classroom teachers of low group students who provide more explicit explanations of how to use these reading skills strategically, would be more effective than teachers who were less explicit in explaining how to use skills.

The subjects were 22 readers randomly assigned to treatment and control groups following a baseline observation. Baseline observation was used to determine each
This study is significant in that a positive correlation was obtained between teacher explanation and student awareness of lesson content. The result suggests that instructional talk has a powerful impact on what students remember and understand.

In their study, (Wittrock & Harold, 1975; Wittrock & Linden, 1981) found that a proactive transfer of learning design was useful in determining a relationship between previous learning and the results of subsequent instruction.

The data from the 336 subjects who completed a total of 100 prior and original learning trials were used to compute the analysis of variance and planned comparisons test.

The results indicated a proactive transfer design was useful in determining a relationship between previous learning and the results of subsequent instruction. Retention was facilitated most when the instructions and feedback were used consistently, and when they were congruent with the student's previously experimentally acquired responses. Retention was the lowest when previously learned responses to sentences were incongruent with the responses reinforced later.

Robinson (1983) indicated that since content is expressed in language, instruction in reading comprehension and studying is, or should be, part of the content curriculum. Furthermore, Paris and his colleagues (Paris,
Gross, & Lipson, 1984) reported that comprehension strategies can result in increased comprehension achievement.

We learn much, if not most, of our vocabulary through direct rewards, rewards for a correct response to words and rewards for our correct use of words.

Brown and Henderson (1978) studied the development of strategies for studying texts. The subjects were 80 college student volunteers paid $2.00 each for their participation.

College students were asked to read stories thoroughly and then to divide the text into idea units. After division into independent units, each story was retyped with one unit per line, and a second group of college students was asked to rate the importance of each unit to the theme of the story using a four-point scale. First, they were asked to eliminate one quarter of the units that they judged to be least important to the theme of the passage. This procedure was then repeated twice more until only one-quarter of the units remained.

The students were tested in small groups or individually, depending on scheduling. All subjects first listened to a tape recording of one of the stories (stories counter-balanced across treatment groups), while they simultaneously read a printed version. They were randomly assigned to four treatment groups. Half the students received an immediate test as soon as the written version of the story was removed. The remainder were permitted five
minutes interaction with the story prior to their recall attempt.

The students were subdivided into those receiving incidental and those receiving intentional instructions. Prior to hearing a tape recording of the story, the incidental group was told that they were collecting foreign folk tales that illustrated traditional morals. They intended to use the stories to study moral development in children cross-culturally. They were to listen to the story and then answer a questionnaire concerning the moral of the story. As soon as the story ended one group were asked to recall the gist in their own words; another group were given the stories to consider and asked to write a brief commentary on the moral and the suitability of the story for students.

The intentional group received explicit instructions that they must attempt recall. The immediate group was tested for recall after hearing the story with no chance to study; the delayed group received five minutes of extra study with the written passage and were told to do anything they wanted to do in order to improve recall.

The results revealed no differences attributable to sex of subject or to study, the intentional group was better able to make use of the extended interaction with the story than was the incidental group. The intentional group was better able to make use of the extended interaction with the story than was the incidental group. A2 (intentional-incidental)
mixed analysis of variance revealed a significant main effect of intentionality.

The delay group outperformed the immediate group only in a significant main effect for immediate delay. The improvement in the delay group was entirely on the two highest levels of importance.

Kitao (1982) reported the common problems instructors encounter in helping students improve their English language reading skills and offered practical exercises designed to enhance the reading experience for foreign students. Following an introductory section that discussed the particular difficulties reading in English presents for Japanese students, the paper discussed the selection of materials and urged the teacher to keep in mind the students' purposes and interests in studying English, the diverse literacy forms that can be used, as well as other considerations, such as bias, length, difficulty, and the advisability of using reading materials written in standard formats. The paper subsequently enumerated the basic skills necessary for specific information, i.e., recognizing the main idea of selection, noting details and understanding relationships, and understanding the sequence of ideas or events. The major portion of the article discussed the purpose and development of various types of reading exercises, such as (a) skimming exercises, (b) comprehension
questions, (c) chart exercises, (d) organization exercises, (e) visualization exercises, (f) close exercises, (g) outline exercises, and (h) discussion questions.

Bruder and Henderson (1985) stated that an effective reading course in English for academic purposes must be based on an understanding of the reading process (including perceptual skills and cognitive processing), an understanding of problems in mastering the English writing system (punctuation, decoding handwriting, and confusion between typefaces), and an awareness of the deficiencies the particular learner group must overcome. Assessment of the learner's reading skills by a variety of techniques, in the context of their language backgrounds, can then lead to development of materials and methods for effective reading instruction. English as a second language programs have adopted a number of approaches to these issues; both theories and procedures are in the process of being refined.

Jenkins, Stein, and Wysocki (1984) stated that new vocabulary knowledge can be acquired through incidental learning of word meanings from context. They arranged relatively normal reading conditions into which they introduced unfamiliar words in paragraphs. Each paragraph was written around a central concept that incorporated the unfamiliar word (e.g., an altercation, a provocative event, incarcerating someone) so that the unfamiliar word occupied a conceptually relevant role in the paragraph.
One hundred and twelve students from four classrooms participated in this study. The 18 vocabulary words selected were designated as low frequency words for students. For each word, 10 paragraphs were written, each approximately 4 to 6 sentences long. The target words remained the same part of speech throughout all paragraphs, although minor changes in tense and number were permitted; they appeared only once in each paragraph, the target word appeared before its implied meaning. To ensure that the students read the paragraphs, two or three literal recall questions were placed at the end of each paragraph.

Four posttest measures were administered: Three vocabulary tests and one reading comprehension test. Of the three vocabulary tests, two were definitional and one was sentence completion. Each vocabulary posttest contained written directions and a sample test item that was completed by the whole class as a group. In addition, one passage (approximately 200 words) was written for each word set and contained all six words in each of the three sets. A set of seven comprehension questions was constructed for each of the three reading passages.

On the first treatment day, teachers told their classes they would be participating in extra reading practice for the next 10 days. Students were also informed that on some days they would receive reading material and on other days they would have to furnish their own. The teacher distributed a
folder containing that day's materials and instructed the students to examine the first page which contained the vocabulary definitions and example sentences for the preexposure condition after allowing students two minutes to study this sheet.

Next, students were instructed to read the paragraphs and answer the questions at the end of each paragraph. Twenty minutes daily were allowed for completing the paragraph/question worksheets. The teacher answered questions about decoding individual words, but gave no help with vocabulary or comprehension questions.

The same procedure was followed for 10 days except that the preexposure sheet did not appear after the first day. Students in the 10 context presentations (paragraphs) read 6 new paragraphs each day, one for each target word. Students in the 6-context presentations read 6 paragraphs on day 1, 3, 5, 7, 9, and 10. In the 2-context (number of paragraphs presented), students read 6 paragraphs on day 1 and 10. In this manner, recency effects for preexposure and context presentations were controlled across context levels.

Posttests, administered by the classroom teachers, began 2 days after the 10 day treatment phase. On the vocabulary testing, students were given two reading comprehension tests, an experimental passage (containing target words that they had read in the paragraph tasks) and a control passage (containing target words that they had not read). After the
first passage was read it was collected and students answered seven comprehension questions. Then the next passage was given, followed by questions. Order of experimental and control passages was counterbalanced.

The results were separate four-factor analyses of variance consisting of 2 (ability levels) x 2 (prior exposure) x 3 (context presentations) x 3 (word sets), with repeated measures on the second factor, were performed on each of the vocabulary knowledge measures. All main effects and interactions were tested, only effects that were consistent on at least two of the three vocabulary measures are reported.

Three findings were consistent across all vocabulary learning measures. First, the word set factor was significant: for supply definition $F(2,89) = 5.9$; for select definition $F(2,88) = 15.1$; and for sentence completion $F(2,89) = 24$ (all p's <.001). Second higher ability students outperformed lower ability students with means of 57% and 29% $F(1,89) = 41.1$ for supply definition; 75% and 51%, $F(1,88) = 30$ for select definition; and 78% and 83%, $F(1,89) = 32.7$ for sentence completion (all p's <.001). A third consistency across all vocabulary measures was a main effect for prior exposure, with means of 54% and 33%, $F(1,89) = 26.7$ for supply definition; 69% and 57%, $F(1,88) = 14.4$ for select definition; and 70% and 58%, $F(1,89) = 13.1$ for sentence completion (all p's <.001).
Martin, McChesney, Whalley, and Devlin (1977) reported that vocabulary learning is a continuous process for all educated people. There are hundreds of thousands of English words. A foreign student cannot expect to learn all of them within a few years. Clearly, foreign students must be selective, must determine which will be the most useful of the words they encounter; that is, whether a word is useful for their passive reading vocabulary or for their active speaking and writing vocabulary.

Statement of the Problem

The purpose of this study was to develop and examine a program to increase the language proficiency of international students. More specifically, the focus of the study was to help new international students from different cultures improve their reading, writing, and other language functions.

Research Questions

Potentially, through a systematic program, reading and writing can be improved by increasing speed and accuracy, vocabulary and comprehension.

This study attempts to answer the following questions through the data obtained from new international students attending Western Michigan University.

1. What are the differences in the speed and accuracy, the vocabulary, and the comprehension performance of
international students attending Western Michigan University at the beginning and at the end of a systematic program?

2. What is the difference in the speed and accuracy, vocabulary, and comprehension performance of international students attending Western Michigan University who receive a vocabulary and writing treatment package and those who do not?
CHAPTER II

DESIGN AND METHOD

The purpose of this study was to develop and examine a systematic program to increase the language proficiency of new international students.

Research

Setting

All procedures were conducted at Western Michigan University with the international students enrolled in a course entitled, "Effective College Reading for International Students". This course focused on vocabulary development and comprehension, including an emphasis on correct pronunciation, word analysis, factual and inferential thinking, and synthesis of ideas. The experiment was conducted in a normal classroom furnished with chairs, tables, and a blackboard.

Subjects

The population in this study was 44 international students. This represents all students in course during the semester.
For purpose of analysis, the students who participated in the fall semester (n=44) were divided into two groups: the control group and the experimental group. The students were selected as to which class they were in by the department scheduler. The control group consisted of 21 international students, 10 females and 11 males. Their native languages were Spanish, Arabic, Chinese, Japanese, and Malaysian.

The experimental group consisted of 23 international students, 17 females and 6 males. Their native languages were Spanish, Chinese, Japanese, and Hebrew.

Apparatus and Materials

The course was designed to provide the student with skills in vocabulary development, comprehension, writing, and reading efficiency. Attention was given to the effective use of text and reference books in academic subjects, inferential reading, and synthesis of main ideas.

The students were expected to learn how to use the various reading skills needed to perform effectively in the courses they were taking. This is reading as a process for becoming a more skillful reader.

Methodology

Both sections were given the Gates and MacGinitie Test (1978) as pretest. The Gates and MacGinitie Reading Survey
(E) Form 1 (1978) was also administered during the first week of the fall semester, 1987. These tests are based on a new, nationwide standardization, and have been validated with students born in the United States. No literature was found reporting use with international (foreign) students.

At the end of the semester, all sections were given the posttest, Gates and MacGinitie Reading Test Survey (E) Form 2 (1978) which was administered in the classroom during the last week of the semester.

The sessions in this study were run from September to December, 1987. Students were divided into two groups: the control group and the experimental group. The students were randomly selected as to which class they were in by the department individual in charge of class scheduling.

Both sections took the pretest at the beginning of fall semester. The experimenter worked just twice with the control group in the classroom during the first week and the last week of the semester for pre and posttest. Over the next 12 weeks a program was implemented with the experimental group to improve reading and writing. One hundred and twenty vocabulary words were selected from the Frequently Misspelled Words list (see Appendix I).

Ten vocabulary words were given to the students in the experimental group only on a weekly basis. They were asked to write and create a story using these vocabulary words.
The experimenter made an analysis of the written stories by using the readability graph developed by Fry (1969, 1977). The graph (see Appendix A) was designed to identify the grade-level score for difficulty of written materials. Attendance was taken and verbal feedback was given by the experimenter about two variables that were used: sentence length and word length.

Assignment of the next 10 vocabulary words occurred so that the students would write another story for the next class and so on.

**Experimental Design**

The basic design is a pretest-posttest comparison between an experimental group and a control group. Both groups received the standard treatment, a course taught to international students.

A treatment package was applied to the experimental group. In addition, a within subjects design was used to assess the effectiveness of the treatment package.

T-tests were used to determine the difference between pretest and posttest scores. Hence, data were entered in a difference array. Each treatment was considered a variable and entered in a different column. We were interested in differences between means; was there a real difference between the pretest and posttest means? (Was the standard treatment effective?) When we were comparing the mean of the
experimental group versus the mean of the control group, we were interested in whether the experimental treatment had any effect.

A letter granting approval (see Appendix J) from the Department of Education and Professional Development, Western Michigan University, and approval from the Human Subjects Institutional Review Board (see Appendix K) were obtained to conduct this study.
CHAPTER III

RESULTS

The primary purpose of the study was to develop a program to increase the language proficiency of new international students. The Gates and MacGinitie Reading Tests (1978) were used in assessing language proficiency. Speed and accuracy, vocabulary, and comprehension scores were obtained for experimental and control groups. The students in the experimental group were given 10 vocabulary words weekly and they were asked to create a story using these vocabulary words. An analysis of the stories using the Fry formula was used to identify the grade-level score for material written each week.

This chapter presents the findings of the study. The first section reports the findings that resulted from testing the hypotheses to determine if a significant difference existed in the mean scores of the experimental group against the mean scores of the control group. The t-test for independent means with an alpha level of .05 was used. The second section reports findings regarding the difference in the speed and accuracy, vocabulary, and comprehension performances of international students who received a vocabulary and writing treatment package and those who did
not. The third section reports the findings of secondary analysis of the data, done to determine whether or not the gain in scores in the experimental group was a direct result of teaching vocabulary words that appeared on the test.

Testing the Hypotheses

This section presents the findings resulting from testing the hypotheses. The t-test for independent means was used to compare the mean pretest and posttest scores of both groups and to compare mean scores of the control and experimental groups. A .05 level of significance was selected as the criterion at which the null hypothesis could be rejected.

Hypothesis One

Part of the first research hypothesis stated that there are significant differences in the speed and accuracy of performance of international students attending Western Michigan University from the beginning to the end of the program.

There were 36 test items related to speed and accuracy with 5 answering options for each item (A, B, C, D, and E). Data in Table 1 reveal that the speed and accuracy posttest scores (mean = 26.88; standard deviation = 6.21) for the experimental group tended to be higher than the speed and accuracy pretest scores (mean = 8.83; standard deviation =
2.21) for that group.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference between the beginning and the end of the program. Statistical analysis using a t-test revealed a 13.10 t value and 0.000 probability that showed significant differences in speed and accuracy. Using the .05 criterion of rejection, the null hypothesis could be rejected.

Table 1

A Comparison of the Mean Scores for the Experimental Group in Speed and Accuracy for the Pretest and Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed and Accuracy - Pretest</td>
<td>23</td>
<td>8.83</td>
<td>2.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed and Accuracy - Posttest</td>
<td>23</td>
<td>26.83</td>
<td>6.21</td>
<td>13.10</td>
<td>0.000</td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

Data in Table 2 reveal that the speed and accuracy pretest score (mean = 7.52; standard deviation = 1.66) for the control group was less than the speed and accuracy posttest score (mean = 8.62; standard deviation = 1.75) for that group.
The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference. Statistical analysis using a t-test revealed a 2.08 t value and with a probability of 0.044 which showed significant differences in speed and accuracy for the control group. Using the .05 level of significance as criterion for rejection, the null hypothesis could be rejected.

Table 2

A Comparison of the Mean Scores for the Control Group in Speed and Accuracy for the Pretest and Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed and Accuracy - Pretest</td>
<td>21</td>
<td>7.52</td>
<td>1.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed and Accuracy - Posttest</td>
<td>21</td>
<td>8.62</td>
<td>1.75</td>
<td>2.08</td>
<td>0.044</td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

A second part of the first hypothesis stated that there would be a significant difference in the vocabulary performance of international students attending Western Michigan University between the beginning and the end of the program.
There were 50 test items related to the vocabulary with 5 answering options for each item. Data in Table 3 reveal that vocabulary pretest scores (mean = 21.43; standard deviation = 3.44) for the experimental group tended to be less than the vocabulary posttest scores (mean = 40.74; standard deviation = 3.99) for that group.

The hypothesis was tested using a $t$-test for independent means with an alpha level of .05 to determine the significant difference between the beginning and the end of the program. Statistical analysis using a $t$-test revealed a $t$ value of 17.58 with a probability of 0.000 which showed significant differences in vocabulary test scores. Using the .05 criterion of rejection, the null hypothesis could be rejected.

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>$t$ Value</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary -</td>
<td>23</td>
<td>21.43</td>
<td>3.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
<td>17.58</td>
<td>0.000</td>
</tr>
<tr>
<td>Vocabulary -</td>
<td>23</td>
<td>40.74</td>
<td>3.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.
Data in Table 4 reveal that vocabulary pretest scores (mean = 19.86; standard deviation = 4.79) for the control group tended to be less than the vocabulary posttest scores (mean = 21.67; standard deviation = 5.50) for that group.

The hypothesis was tested using a t-test for independent means with an alpha of .05 to determine the significant difference at the beginning and at the end of the program. Statistical analysis using a t-test revealed a t value of 1.14 with a probability of 0.26. Using the .05 level of significance as criterion for rejection, the null hypothesis could not be rejected.

Table 4
A Comparison of the Mean Scores for the Control Group in Vocabulary for the Pretest and Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary -</td>
<td>21</td>
<td>19.86</td>
<td>4.79</td>
<td>1.14</td>
<td>0.26</td>
</tr>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary -</td>
<td>21</td>
<td>21.67</td>
<td>5.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

Finally, the first hypothesis stated that there would be a significant difference in the comprehension performance of international students attending Western Michigan University at the beginning and end of the program.
There were 52 test items related to comprehension with 5 answering options for each item.

Data in Table 5 reveal that comprehension pretest scores (mean = 25.17; standard deviation = 5.64) for the experimental group tended to be less than the comprehension posttest scores (mean = 44.61; standard deviation = 4.06) for that group.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference at the beginning and at the end of the program. Statistical analysis using a t-test revealed a t value of 13.41 with a probability of 0.000. Using the .05 level of significance as criterion for rejection, the null hypothesis could be rejected.

Table 5

A Comparison of the Mean Scores for the Experimental Group on Comprehension for the Pretest and Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension -</td>
<td>23</td>
<td>25.17</td>
<td>5.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
<td>13.41</td>
<td>0.000</td>
</tr>
<tr>
<td>Comprehension -</td>
<td>23</td>
<td>44.61</td>
<td>4.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.
Data in Table 6 reveal that comprehension pretest scores (mean = 23.24; standard deviation = 6.39) for the control group tended to be less than the comprehension posttest scores (mean = 25.81; standard deviation = 5.97) for that group.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference at the beginning and at the end of the program. Statistical analysis using a t-test revealed a t value of 1.35 with a probability of 0.19. Using the .05 level of significance as criterion for rejection, the null hypothesis could not be rejected.

Table 6

A Comparison of the Mean Scores for the Control Group on Comprehension for the Pretest and Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension -</td>
<td>21</td>
<td>23.24</td>
<td>6.39</td>
<td>1.35</td>
<td>0.19</td>
</tr>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension -</td>
<td>21</td>
<td>25.81</td>
<td>5.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.
Hypothesis Two

The second research hypothesis stated that there will be a difference in the writing performance of international students (attending Western Michigan University) from the beginning to the end of the treatment program.

Over the twelve weeks, a program was designed to improve writing performance. Ten vocabulary words were given to the students in the experimental group only. They were asked to write and create a story using these vocabulary words. The following day the experimenter made an analysis of writing stories using the readability graph developed by Fry (1977).

Data in Appendix E reveal that during the treatment program, the experimental group improved writing performance. During the first week, the mean was 38.88 with a standard deviation of 12.07. The minimum score for all the experimental group was 16.66% while the maximum score was 55.55%. At the end of the program the mean was 70.53 with a standard deviation of 15.66. The minimum score for all the experimental group was 38.88% and the maximum score was 94%.

The findings of the data analysis with regard to each week show at the end of the second week, a mean of 37.68 which was less than the first week and a standard deviation of 8.53, while the minimum score was 22.22% (an increase over the first week), but a maximum score of 50% which was less than the first week.
The experimental group's treatment scores at the end of the third week had a mean of 38.64 which was similar to the first week and greater than the second week. The standard deviation was 11.41 with a minimum score of 16.66% and a maximum score of 55.55%; these were the same as the first week but greater than the second week.

The experimental group treatment scores at the end of the fourth week (mean of 43.71) tended to be greater than the first three weeks, with a standard deviation of 8.9, a minimum score of 22.22% and a maximum of 55.55% similar to the first and third weeks.

The treatment scores for the fifth week (mean of 45.65) tended to be greater than the first four weeks, with a standard deviation of 14.50, a minimum score of 22.22% and a maximum score of 72.22% which was greater showing some improvement.

The sixth week (mean of 46.85) tended to be greater than the first five weeks with a standard deviation of 11.46 and the same minimum score (22.22%). The maximum score (61.11%) tended to be less than Week 5 but greater than the first four weeks.

The treatment scores at the end of Week 7 (mean 50.72) tended to be greater than the weeks before, with a standard deviation of 12.79, a minimum score of 27.77% showing an increase over the weeks before. The maximum score (66.66%) was less than Week 5 but greater than other weeks before.
Week 8 (mean of 55.79) was greater than the weeks before, with a standard deviation of 15.30, a minimum score of 22.22% which was less than Week 7 but the same as Weeks 2, 4, 5, and 6. The maximum score (83.33%) tended to be an improvement in their performance for the weeks before.

The experimental group treatment scores for week 9 (mean of 63.04) was greater, with a standard deviation of 15.49, a minimum score of 27.77% which tended to be the same as week 7, but greater than other weeks. The maximum score (94.4%) was greater than other weeks and displayed a great improvement in writing skills and an increase in performance.

Week 10 with a mean of 65.94 which was greater, had a standard deviation of 17.75, a minimum score of 22.22% (the weekly average), and a maximum score of 94.4% which was the same as Week 9 but greater than the other weeks.

For Week 11 there was a mean of 73% which was greater, and a standard deviation of 17.54, a minimum score of 38.88% which tended to be greater than other weeks, and a maximum score of 94.4% which was the same as Weeks 9 and 10.

Week 12 with a mean of 70.53 was less than Week 11 but greater than the other weeks with a standard deviation of 15.66, a minimum score of 38.88% and a maximum score of 94.4% which was the same as Week 11 and greater than other weeks.

In conclusion, there were significance differences in writing performance of international students at the beginning and the end of the program.
There were 36 test items related to vocabulary with five answering options for each item. Data in Table 8 reveal that vocabulary pretest scores for the experimental group (mean = 21.43) were higher than the control group (mean = 19.86). The standard deviation for the experimental group was 3.44 while the standard deviation for the control group was 4.79.

The hypothesis was tested using a $t$-test for independent means with an alpha level of .05 to determine the significant difference in the vocabulary pretest of the experimental and control groups. Statistical analysis using a $t$-test reveals a $t$ value of 1.25 with a probability of 0.22. Using the .05 level of significance as criterion for rejection, the null hypothesis for experimental and control group vocabulary tests could not be rejected.

Data in Table 8 reveal that the experimental group vocabulary scores for the posttest (mean = 40.74) were higher than the control group (mean = 21.67). The standard deviation for the experimental group was 3.99 while the standard deviation for the control group was 5.50.
Table 7
A Comparison of the Mean Scores for the Experimental and Control Groups for Vocabulary on the Pretest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t  Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>23</td>
<td>21.43</td>
<td>3.44</td>
<td>1.25</td>
<td>0.22</td>
</tr>
<tr>
<td>Control Group</td>
<td>21</td>
<td>19.86</td>
<td>4.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference in vocabulary posttest scores for both the experimental and control groups. Statistical analysis using a t-test revealed a t value of 13.06 with a probability of 0.000. Using the .05 level of significance as criterion for rejection, the null hypothesis the difference between the experimental and control groups for the vocabulary test could be rejected. There were 50 test items related to speed and accuracy with five answering options for each item.
Table 8
A Comparison of the Mean Scores for the Experimental and Control Groups for Vocabulary on the Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>23</td>
<td>40.74</td>
<td>3.99</td>
<td>13.06</td>
<td>0.000</td>
</tr>
<tr>
<td>Control Group</td>
<td>21</td>
<td>21.67</td>
<td>5.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

Data in Table 9 reveal that the speed and accuracy pretest scores for the experimental group (mean = 8.83) were higher than the scores for the control group (mean = 7.52). The standard deviation for the experimental group was 2.21 while the standard deviation for the control group was 1.66.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the independent means difference in speed and accuracy on the pretest for both the experimental and control groups. Statistical analysis using a t-test revealed a t value of 2.22 with a probability of 0.32. Using the .05 level of significance as criterion for rejection, the null hypothesis of no difference between experimental and control groups in speed and accuracy could not be rejected.
Table 9
A Comparison of the Mean Scores for the Experimental and Control Groups for Speed and Accuracy on the Pretest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>23</td>
<td>8.83</td>
<td>2.21</td>
<td>2.22</td>
<td>0.32</td>
</tr>
<tr>
<td>Control Group</td>
<td>21</td>
<td>7.52</td>
<td>1.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

Data in Table 10 revealed that the scores for the experimental group on the speed and accuracy posttest (mean = 26.83) was higher than the control group (mean = 8.62). The standard deviation for the experimental group was 6.21, while the standard deviation for the control group was 1.75.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference in speed and accuracy scores between the experimental and control groups. Statistical analysis using a t-test revealed a t-value of 13.50 with a probability of 0.000. Using the .05 level of significance as criterion for rejection, the null hypothesis of no difference between the experimental and control groups in speed and accuracy on the posttest could be rejected.
Table 10
A Comparison of the Mean Scores for the Experimental and Control Groups for Speed and Accuracy on the Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>23</td>
<td>26.83</td>
<td>6.21</td>
<td>13.50</td>
<td>0.000</td>
</tr>
<tr>
<td>Control Group</td>
<td>21</td>
<td>8.62</td>
<td>1.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

There were 52 test items related to comprehension with five answering options for each item.

Data in Table 11 revealed that the experimental group comprehension pretest scores (mean = 25.17) were higher than the scores for the control group (mean = 23.24). The standard deviation for the experimental group was 5.64, while the standard deviation for the control group was 6.39.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference of the comprehension pretest for both the experimental and control groups. Statistical
analysis using a t-test revealed a t value of 1.06 with a probability of 0.30. Using the .05 level of significance as criterion for rejection, the null hypothesis for a difference between the experimental and control groups on the comprehension test could not be rejected.

Table 11
A Comparison of the Mean Scores for the Experimental and Control Groups for Comprehension on the Pretest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>23</td>
<td>25.17</td>
<td>5.64</td>
<td>1.06</td>
<td>0.30</td>
</tr>
<tr>
<td>Control Group</td>
<td>21</td>
<td>23.24</td>
<td>6.39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

Data in Table 12 revealed that the experimental group scores on the comprehension posttest (mean = 44.61) were higher than those for the control group (mean = 25.81). The standard deviation for the experimental group was 4.06, while the standard deviation for the control group was 5.97.

The hypothesis was tested using a t-test for independent means with an alpha level of .05 to determine the significant difference in comprehension posttest scores
for both the experimental and control groups. Statistical analysis using a t-test revealed a t value of 12.09 with a probability of 0.000. Using the .05 level of significance as criterion for rejection, the null hypothesis of no difference between the experimental and control groups on the comprehension posttest could be rejected.

Table 12
A Comparison of the Mean Scores for the Experimental and Control Groups for Comprehension on the Posttest

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>t Value</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td>23</td>
<td>44.61</td>
<td>4.06</td>
<td>12.09</td>
<td>0.000</td>
</tr>
<tr>
<td>Control Group</td>
<td>21</td>
<td>25.81</td>
<td>5.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the .05 level.

Findings of Secondary Analysis Data

Secondary analyses of the data were performed to aid in interpreting the results. Why did the practice in writing and using new vocabulary items each week produce gains in the scores? One possibility is that the vocabulary words used were also words tested on the Gates and MacGinitie tests (1978).
All the vocabulary words from the List of Frequently Misspelled Words were checked to see if they were words used in the test. Test items containing a word from the list are counted as "vocabulary questions" in Table 13, Table 14 and Table 15. "Non-vocabulary questions" refer to test items not containing a vocabulary item assigned.

Data in Table 13 show a pretest-posttest gain in speed and accuracy of 160% for questions containing assigned vocabulary words, and 200% for nonvocabulary questions in the experimental group. This was higher than the pretest-posttest gain in speed and accuracy of 78% for questions containing assigned vocabulary words, and 4% for nonvocabulary questions in the control group.

As might be expected, the gain in vocabulary was greatest for items containing the words assigned during the course: 130% for vocabulary questions compared to a gain of 83% for nonvocabulary questions in the experimental group. The figure also shows a higher gain for the experimental group. The control group gain for vocabulary questions was 2% compared to gain of 7% for nonvocabulary questions.

The pretest-posttest gain in comprehension (48%) for vocabulary questions in the experimental group shows less than the control group (55%). However, the experimental group shows higher percentage gain overall than the control group (see Figures 1 and 2). The pretest-posttest gain in comprehension (67%) for nonvocabulary questions in the
experimental group, and 25% in the control group.

Data in Table 14 show a speed and accuracy of 30% for vocabulary questions and 24% for nonvocabulary questions in the experimental group. The accuracy for vocabulary questions was 40% compared to an accuracy of 43% for nonvocabulary questions. The comprehension accuracy for vocabulary questions was 50% compared to 46% for nonvocabulary questions in the experimental group (see Figure 3).

The posttest accuracy in speed and accuracy for vocabulary questions, 76%, was higher than the accuracy of 73% for nonvocabulary questions. The accuracy in vocabulary for vocabulary questions was 93% compared to 79% for nonvocabulary questions. The accuracy of comprehension for vocabulary questions was 74% compared to 77% for nonvocabulary questions in the experimental group (see Figure 4).

Data in Table 15 show a posttest gain in speed and accuracy of (25%) for vocabulary questions was higher than the pretest in speed and accuracy of (14%) for vocabulary questions in the control group. The posttest gain in vocabulary of (35%) for vocabulary questions was higher than the pretest vocabulary (34%) for vocabulary questions in control group. However, the posttest gain in comprehension (45%) for vocabulary questions was higher than the pretest in comprehension (29%) for vocabulary questions in control
group (see Figure 5).

The posttest gain in speed and accuracy of (24%) for nonvocabulary questions was higher than pretest in speed and accuracy of (23%) for nonvocabulary questions in the control group. The posttest gain in vocabulary of (44%) for nonvocabulary questions was higher than pretest in vocabulary of (41%) for nonvocabulary questions in the control group. However, the posttest gain in comprehension of (50%) for nonvocabulary group was higher than pretest in comprehension of (40%) for nonvocabulary questions in control group (see Figure 6).

These analyses show that the gains were not a simple result of teaching the test. The experimental group showed higher gains than the control group on both vocabulary and nonvocabulary questions.

Moreover, the experimental group gains in speed and accuracy and in comprehension were about the same (76% and 73% for speed and accuracy; 74% and 77% for comprehension) for vocabulary and nonvocabulary questions. Only the gains in vocabulary (93% and 79%) show a difference that might be attributed to teaching the test; however, the number is rather small for vocabulary questions on the vocabulary posttest so we should place little confidence in the 93% versus the 79% difference.
### Table 13
The Percentage of Posttest Minus Pretest Divided by Pretest for Experimental Group

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Speed &amp; Accuracy</th>
<th>Vocabulary</th>
<th>Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary Questions</td>
<td>160%</td>
<td>130%</td>
<td>48%</td>
</tr>
<tr>
<td>Nonvocabulary Questions</td>
<td>200%</td>
<td>83%</td>
<td>67%</td>
</tr>
</tbody>
</table>

| The Percentage of Posttest Minus Pretest Divided by Pretest for Control Group |
|------------------|------------|------------|
| Control Group | Speed & Accuracy | Vocabulary | Comprehension |
| Vocabulary Questions | 78% | 2% | 55% |
| Nonvocabulary Questions | 4% | 7% | 25% |
Figure 1. Data From "Experimental Group".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 2. Data From "Control Group".
Table 14
The Percentage of Items Correct Divided by Total Items for Experimental Group

<table>
<thead>
<tr>
<th></th>
<th>Pretest (Vocabulary Questions)</th>
<th>Posttest (Vocabulary Questions)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>49</td>
<td>102</td>
<td>105</td>
</tr>
<tr>
<td>Wrong</td>
<td>112</td>
<td>151</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>161</td>
<td>253</td>
<td>207</td>
</tr>
<tr>
<td>%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Pretest (Nonvocabulary Questions)</th>
<th>Posttest (Nonvocabulary Questions)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>162</td>
<td>391</td>
<td>460</td>
</tr>
<tr>
<td>Wrong</td>
<td>505</td>
<td>506</td>
<td>529</td>
</tr>
<tr>
<td>Total</td>
<td>667</td>
<td>897</td>
<td>989</td>
</tr>
<tr>
<td>%</td>
<td>24%</td>
<td>43%</td>
<td>46%</td>
</tr>
</tbody>
</table>
Figure 3. Data From "Experimental Group".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 4. Data From "Experimental Group".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 15

The Percentage of Items Correct Divided by Total Items for Control Group

<table>
<thead>
<tr>
<th>Pretest (Vocabulary Questions)</th>
<th>Speed &amp; Accuracy</th>
<th>Vocab. Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>18</td>
<td>80</td>
</tr>
<tr>
<td>Wrong</td>
<td>129</td>
<td>151</td>
</tr>
<tr>
<td>Total</td>
<td>147</td>
<td>231</td>
</tr>
<tr>
<td>%</td>
<td>14%</td>
<td>34%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Posttest (Vocabulary Questions)</th>
<th>Speed &amp; Accuracy</th>
<th>Vocab. Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>54</td>
<td>15</td>
</tr>
<tr>
<td>Wrong</td>
<td>156</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>210</td>
<td>42</td>
</tr>
<tr>
<td>%</td>
<td>25%</td>
<td>35%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pretest (Nonvocabulary Questions)</th>
<th>Speed &amp; Accuracy</th>
<th>Vocab. Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>140</td>
<td>342</td>
</tr>
<tr>
<td>Wrong</td>
<td>469</td>
<td>477</td>
</tr>
<tr>
<td>Total</td>
<td>609</td>
<td>819</td>
</tr>
<tr>
<td>%</td>
<td>23%</td>
<td>41%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Posttest (Nonvocabulary Questions)</th>
<th>Speed &amp; Accuracy</th>
<th>Vocab. Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>132</td>
<td>440</td>
</tr>
<tr>
<td>Wrong</td>
<td>414</td>
<td>568</td>
</tr>
<tr>
<td>Total</td>
<td>546</td>
<td>1008</td>
</tr>
<tr>
<td>%</td>
<td>24%</td>
<td>44%</td>
</tr>
</tbody>
</table>
Figure 5. Data from "Control Group".
POST TEST

Figure 6. Data From "Control Group".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CHAPTER IV

DISCUSSION AND RECOMMENDATIONS

Discussion

The purpose of this study was to develop a systematic program to increase the language proficiency of international students. The students improved from the beginning to the end of the treatment package, and there was a significant difference between the experimental group and the control group at the end of the program. Although both groups improved, the experimental group improved more than the control group in all three components: speed and accuracy, vocabulary, and comprehension.

The experimental group obtained a higher mean score than the control group at the beginning of the treatment program. This difference was not statistically significant. However, both groups had improved at the end of the program, while the experimental group was statistically significant than the control group. That meant the treatment package was effective.

The results from secondary analyses of the data were performed to help in interpreting the findings. The experimental group's percentage (vocabulary and nonvocabulary
questions) for the two components, speed and accuracy, and vocabulary was higher than that of the control group. The most important finding of the study, however, was that the control group (vocabulary questions) had a higher percentage than the experimental group in the comprehension component despite the fact that the experimental group achieved higher grades in both the pretest and posttest than the control group.

Data show the posttest gain in speed and accuracy, vocabulary, and comprehension for vocabulary and nonvocabulary questions had a higher gain than pretest in the experimental group. The posttest gain in speed and accuracy, vocabulary, and comprehension for vocabulary and nonvocabulary questions had a higher gain than the pretest in control group. However, the experimental group showed higher gain than the control group on both vocabulary and nonvocabulary questions.

Data shows speed and accuracy, vocabulary, and comprehension in pretest-posttest for vocabulary and nonvocabulary for the experimental group was higher than the control group.

The data show that the experimental group scored slightly higher on the pretest than the control group, although the differences were not statistically significant. The experimental group's superiority might be because the researcher administered the pre and posttest for both groups.
and may have "unconsciously" demanded less from the control group. Alternatively, the students in the control group may have had a worse daily schedule that conflicted with their performance, or the students in the experimental group might have been superior.

Recommendations

Based on the findings and conclusions of this study, the following recommendations are presented for planning, developing, and conducting a program for new international students, as well as for future research.

1. New international students should write all vocabulary words listed weekly through the treatment package on a packet of 3X5 inch index cards. These cards should be used and blank cards should be carried by the student every day. When they hear or read a word or phrase that they need to learn, they should use these vocabulary words when they write down a story for the following week.

2. Future research studies should attempt to identify more vocabulary items than the ones used each week from the List of Frequently Misspelled Words. The vocabulary words should contain more vocabulary questions than the nonvocabulary questions covered in the required book and test items. This study indicated that the gain in vocabulary questions was statistically more significant than the nonvocabulary questions.
3. Future research studies should strive for a much larger sample. Similar group research may produce different findings by comparison between experimental and control groups in the fall and winter semester.
REFERENCES


Blair, L. (1983). ESL students and writing acquisition, George Mason University, VA.

discrimination and sentence imitation in children. Child
Development, 46, 444-451.

to learn: On training students to learn from texts.
Evaluation Researcher, 10, 14-21.

Bumer, J. (1966). Toward a theory of instruction. Boston,
MA: Harvard University Press.

The Eastern Michigan University Conference on Languages for
Business and the Professions, 4th, Dearborn, MI, May 2-4.

Centra, J. (1977). Student ratings of instruction and their
relationship to student learning. American Education

from 5 to 10. Cambridge, MA: Massachusetts Institute of
Technology Press.

Choy, S. C., & Davenport, B. M. (1986). The TOEFL:
Incomplete test of English proficiency. College Teaching,

Experimental Education, 36, 14-25.

Council of Graduate Schools in the United States, Washington,
D. C. (1980). The Foreign Student in American Graduate
Schools, 1, 28.


Quarterly, 12(4), 439-448.

DeCharms, R. (1976). Enhancing motivation: Change in the


listening: A comparative study. Journal of Reading
Behavior, 10, 203-207.

Elson, N. (1984). Reading and meaning for university level
ESL students. TESL Talk, 15(3), 26-34.


Shinoda, Y. (Win/Spr., 1983). Some practical advice on teaching English technical writing to foreign students. Technical Writing Teacher, 10 (23), 140-147.


APPENDIX A

Fry Readability Graph
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 outmode or render the earlier (1968) version inaccurate or incorrect; it is an extension. (REPRODUCTION PERMITTED—NO COPYRIGHT)
APPENDIX B

Figures of Weekly Improvement of Performance for all Individuals from the Beginning to the End of Treatment
Figure 7. Performance Scores for All Subjects at Week 1.
Figure 8. Performance Scores for All Subjects at Week 2.
Figure 9. Performance Scores for All Subjects at Week 3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 10. Performance Scores for All Subjects at Week 4.
Figure 11. Performance Scores for All Subjects at Week 5.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 12. Performance Scores for All Subjects at Week 6.
Figure 13. Performance Scores for All Subjects at Week 7.
Figure 16: Performance Scores for All Subjects at Week 10.
Figure 17. Performance Scores for All Subjects at Week 11.
Figure 18. Performance Scores for All Subjects at Week 12.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
APPENDIX C

Figures of Weekly Improvement of Performance for Each Individual From the Beginning to the End of Treatment
WEEKLY

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
APPENDIX D

Figures of Comparison Between Experimental and Control Groups
Data from "CONTROL GROUP"
APPENDIX E

Treatment of the Writing Performance of Experimental Group from the Beginning to the End of the Program
### Appendix E
Treatment of the Writing Performance of Experimental Group from the Beginning to the End of the Program

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum Score</th>
<th>Maximum Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>23</td>
<td>38.88</td>
<td>12.07</td>
<td>16.66</td>
<td>55.55</td>
</tr>
<tr>
<td>Week 2</td>
<td>23</td>
<td>37.68</td>
<td>8.53</td>
<td>22.22</td>
<td>50.00</td>
</tr>
<tr>
<td>Week 3</td>
<td>23</td>
<td>38.64</td>
<td>11.41</td>
<td>16.66</td>
<td>55.55</td>
</tr>
<tr>
<td>Week 4</td>
<td>23</td>
<td>43.71</td>
<td>8.91</td>
<td>22.22</td>
<td>55.55</td>
</tr>
<tr>
<td>Week 5</td>
<td>23</td>
<td>45.65</td>
<td>14.50</td>
<td>22.22</td>
<td>72.22</td>
</tr>
<tr>
<td>Week 6</td>
<td>23</td>
<td>46.85</td>
<td>11.46</td>
<td>22.22</td>
<td>61.11</td>
</tr>
<tr>
<td>Week 7</td>
<td>23</td>
<td>50.72</td>
<td>12.79</td>
<td>27.77</td>
<td>66.66</td>
</tr>
<tr>
<td>Week 8</td>
<td>23</td>
<td>55.79</td>
<td>15.30</td>
<td>22.22</td>
<td>83.33</td>
</tr>
<tr>
<td>Week 9</td>
<td>23</td>
<td>63.04</td>
<td>15.49</td>
<td>27.77</td>
<td>94.44</td>
</tr>
<tr>
<td>Week 10</td>
<td>23</td>
<td>65.94</td>
<td>17.75</td>
<td>22.22</td>
<td>94.44</td>
</tr>
<tr>
<td>Week 11</td>
<td>23</td>
<td>73.18</td>
<td>17.54</td>
<td>38.88</td>
<td>94.44</td>
</tr>
<tr>
<td>Week 12</td>
<td>23</td>
<td>70.53</td>
<td>15.66</td>
<td>38.88</td>
<td>94.44</td>
</tr>
</tbody>
</table>
APPENDIX F

Frequently Misspelled Words
FREQUENTLY MISSPELLED WORDS

<table>
<thead>
<tr>
<th>accidentally</th>
<th>conscience</th>
<th>immediately</th>
<th>psychology</th>
</tr>
</thead>
<tbody>
<tr>
<td>achievement</td>
<td>consensus</td>
<td>indispensable</td>
<td>pursu</td>
</tr>
<tr>
<td>acquaintance</td>
<td>contemptible</td>
<td>irrelevant</td>
<td>questionnaire</td>
</tr>
<tr>
<td>advantageous</td>
<td>convenience</td>
<td>irresistible</td>
<td>receive</td>
</tr>
<tr>
<td>amateur</td>
<td>courageous</td>
<td>knowledge</td>
<td>reference</td>
</tr>
<tr>
<td>analysis</td>
<td>definitely</td>
<td>laboratory</td>
<td>repetable</td>
</tr>
<tr>
<td>anonymous</td>
<td>descent</td>
<td>legible</td>
<td>rhythm</td>
</tr>
<tr>
<td>anxiety</td>
<td>desiring</td>
<td>lieutenant</td>
<td>sarcasm</td>
</tr>
<tr>
<td>applicable</td>
<td>desirable</td>
<td>lightning</td>
<td>scissors</td>
</tr>
<tr>
<td>appreciate</td>
<td>despair</td>
<td>loneliness</td>
<td>secretory</td>
</tr>
<tr>
<td>arctic</td>
<td>disappear</td>
<td>manageable</td>
<td>seize</td>
</tr>
<tr>
<td>argue</td>
<td>discipline</td>
<td>manageable</td>
<td>separate</td>
</tr>
<tr>
<td>auxiliary</td>
<td>efficient</td>
<td>mathematics</td>
<td>serviceable</td>
</tr>
<tr>
<td>awkward</td>
<td>eighth</td>
<td>meant</td>
<td>simulate</td>
</tr>
<tr>
<td>bachelor</td>
<td>either</td>
<td>medieval</td>
<td>sophomore</td>
</tr>
<tr>
<td>baggage</td>
<td>eligible</td>
<td>minimum</td>
<td>souvenir</td>
</tr>
<tr>
<td>beginning</td>
<td>equipped</td>
<td>mortgage</td>
<td>statistics</td>
</tr>
<tr>
<td>believe</td>
<td>exaggerate</td>
<td>necessary</td>
<td>subtle</td>
</tr>
<tr>
<td>benefited</td>
<td>exercise</td>
<td>neither</td>
<td>succeed</td>
</tr>
<tr>
<td>bookkeeper</td>
<td>exhausted</td>
<td>noticeable</td>
<td>successful</td>
</tr>
<tr>
<td>breathe</td>
<td>existence</td>
<td>occasionally</td>
<td>succeed</td>
</tr>
<tr>
<td>bucketfuls</td>
<td>existance</td>
<td>opportunity</td>
<td>surprising</td>
</tr>
<tr>
<td>bureau</td>
<td>foreseeable</td>
<td>Parallel</td>
<td>synonym</td>
</tr>
<tr>
<td>business</td>
<td>fascinating</td>
<td>paralyse</td>
<td>&lt;tablespoonfuls</td>
</tr>
<tr>
<td>calendar</td>
<td>familiar</td>
<td>pastime</td>
<td>thoroughly</td>
</tr>
<tr>
<td>campaign</td>
<td>fiery</td>
<td>plebelan</td>
<td>tragedy</td>
</tr>
<tr>
<td>caricature</td>
<td>financier</td>
<td>possibility</td>
<td>transformable</td>
</tr>
<tr>
<td>catastrophe</td>
<td>foreign</td>
<td>privilege</td>
<td>twelth</td>
</tr>
<tr>
<td>cemetery</td>
<td>forty</td>
<td>procedure</td>
<td>unforgettable</td>
</tr>
<tr>
<td>colonel</td>
<td>government</td>
<td>prominent</td>
<td>vacuum</td>
</tr>
<tr>
<td>coming</td>
<td>guarantee</td>
<td>propaganda</td>
<td>vengeance</td>
</tr>
<tr>
<td>commitment</td>
<td>height</td>
<td>humorous</td>
<td>weird</td>
</tr>
<tr>
<td>committee</td>
<td>heroes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>comparative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>competitive</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G

Approval Letter from the Department of Education and Professional Development
August 21, 1987

Mr. Abatah Daher
1984 Howard St. Apt. 489
Kalamazoo, MI 49008

Dear Mr. Daher:

I have read your dissertation proposal and discussed it with Dr. Lloyd. Dr. Lloyd has agreed to cooperate with you by permitting you to work with students in his ED 106 Reading for International Student course.

Dr. Lloyd and I assume that you and your dissertation chair will assume the responsibility of meeting university requirements for the protection of human subjects.

Best wishes for a successful study.

Sincerely,

Thomas F. Ryan, Chair
Education and Professional Development

cc: Dr. Lloyd - EPD
Dr. Brethower - Psychology

TR/vt
APPENDIX H

Approval Letter from the Human Subjects Institutional Review Board
TO: Abatah Daher, Dale M. Brethower

FROM: Ellen Page-Robin, Chair

RE: Research Protocol

DATE: August 31, 1987

This letter will serve as confirmation that your research protocol, "Developing a systematic program to increase the language proficiency of international students attending Western Michigan University," has been approved as exempt by the HSIRB.

If you have any further questions, please contact me at 383-4917.
BIBLIOGRAPHY


