

© 1979

ANNETTE JOHNSON THOMASON

ALL RIGHTS RESERVED

iversity  
Microfilms  
ernational

N. ZEEB ROAD, ANN ARBOR, MI 48106

THE EFFECT OF WORD FREQUENCY ON COPY DIFFICULTY  
FOR SHORTHAND TESTING MATERIALS

By

Annette Johnson Thomason

Bachelor of Science  
Brigham Young University  
Provo, Utah  
1969

Master of Science  
Brigham Young University  
Provo, Utah  
1971

Submitted to the Faculty of the Graduate College  
of the Oklahoma State University  
in partial fulfillment of the requirements  
for the Degree of  
DOCTOR OF EDUCATION  
May, 1979

Thesis  
1979D  
T463e  
Lop. 2



THE EFFECT OF WORD FREQUENCY ON COPY DIFFICULTY  
FOR SHORTHAND TESTING MATERIALS

Thesis Approved:

*Arnola C. Ownley*  
\_\_\_\_\_  
Thesis Adviser

*Lloyd Wiggins*  
\_\_\_\_\_

*Harold A. Conrad*  
\_\_\_\_\_

*Jeanine H. Rhea*  
\_\_\_\_\_

*Norman N. Durham*  
\_\_\_\_\_  
Dean of the Graduate College

1041537

## ACKNOWLEDGMENTS

Despite the inevitable traumas of a project of this nature, it has been an interesting and rewarding challenge, made so by the encouragement, enthusiasm, and support of the many people who have assisted in so many ways. To them all, I am deeply grateful for their part in making this endeavor an exciting opportunity.

Special thanks and sincere appreciation are especially extended to Dr. Arnola C. Ownby, major adviser, for her generous help, professional example, and expert guidance in the completion of this study.

Sincere appreciation is also extended to the other members of the committee Dr. Jeanine Rhea, Dr. Harold Coonrad, and Dr. Lloyd Wiggins for their continued encouragement and many helpful suggestions.

Special recognition and thanks are given to Dr. David Weeks of the Statistics Department for his invaluable assistance and expertise in the development of the statistical design and the analysis and interpretation of the data.

A large debt of gratitude is due the teachers who assisted in the collection of the data: Mrs. Janice London, Mrs. Shirley Schooley, Dr. Harold Coonrad, Mrs. Denise Hamburger, Mrs. Evelyn Randolph, Mrs. Elise Earl, Mrs. Adelia Farris, Mrs. Helen Nichols, Mrs. Zelma Faulkner, and Mrs. Peggy Butler. Unlimited thanks are also due Miss Denise Flanagan who assisted so ably with the scoring of the many transcripts.

Much gratitude is also expressed for the financial assistance provided by the U. S. Office of Education through the Educational Professions

Development Act and the Vocational Education Graduate Leadership Development Program.

And warmest appreciation and love are extended to the writer's family and friends, especially her mother, Mrs. Ver1 J. Somerville, and sons, Allen and Mark, whose understanding, patience, and sacrifice have made the completion of this study possible.

## TABLE OF CONTENTS

Chapter	Page
I. THE RESEARCH PROBLEM . . . . .	1
Introduction . . . . .	1
Need for the Study . . . . .	5
Purpose of the Study . . . . .	8
Statement of the Problem . . . . .	8
Hypotheses . . . . .	9
Delimitations . . . . .	10
Limitations . . . . .	11
Definition of Terms . . . . .	12
Assumptions . . . . .	13
Summary . . . . .	14
II. RELATED LITERATURE . . . . .	15
Introduction . . . . .	15
Studies Dealing with Copy Difficulty . . . . .	16
Studies Dealing with Related Factors . . . . .	48
Summary . . . . .	49
III. PROCEDURES . . . . .	52
Introduction . . . . .	52
Development of the Test Instrument . . . . .	52
Selection of Participants . . . . .	62
Administration of the Tests . . . . .	65
Evaluation of the Data . . . . .	68
Analysis of the Data . . . . .	69
Summary . . . . .	70
IV. FINDINGS . . . . .	71
Introduction . . . . .	71
Determination of Equivalency of Tests at Same Difficulty Levels . . . . .	72
Comparison of Three Difficulty Levels . . . . .	73
Supporting Evidence . . . . .	78
Summary . . . . .	80

Chapter	Page
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS . . . . .	81
Summary . . . . .	81
Conclusions . . . . .	86
Recommendations . . . . .	87
SELECTED BIBLIOGRAPHY . . . . .	90
APPENDIXES	
APPENDIX A - FORM USED TO RECORD AND TABULATE THE NUMBER OF WORDS NEEDED AND/OR USED IN EACH CATEGORY CONTROLLED . . . . .	97
APPENDIX B - TESTS DEVELOPED FOR THIS STUDY . . . . .	99
APPENDIX C - TRANSCRIPT OF INTRODUCTORY MATERIAL ON THE SIX STUDY TEST TAPES . . . . .	106
APPENDIX D - GUIDELINES FOR TEACHERS . . . . .	108
APPENDIX E - TRANSCRIPT OF INTRODUCTORY MATERIAL ON THE PRELIMINARY TEST TAPES . . . . .	111
APPENDIX F - FORM USED TO RECORD ERROR SCORES . . . . .	113
APPENDIX G - SUMMARY OF RAW SCORES FOR ALL STUDENTS ON SIX STUDY TESTS . . . . .	115
APPENDIX H - SUMMARY OF RAW MEAN TRANSCRIPTION ERROR SCORES FOR EIGHT CLASSES . . . . .	119
APPENDIX I - COMPUTATION OF Z SCORES FOR THREE DIFFI- CULTY LEVEL COMPARISONS . . . . .	121



## LIST OF TABLES

Table	Page
I. Percent of Increase for Words in Each of Three Difficulty Levels . . . . .	55
II. Number of Words and Range Allowed for Brief Forms and Brief Form Derivatives in One- and Three-Minute Segments of the Six Study Tests . . . . .	56
III. Syllabic Intensity for One- and Three-Minute Portions of Six Study Tests . . . . .	57
IV. Number of Different Words for Each of Six Study Tests . .	59
V. Number of Words Allowed in Each Word Group for Three Difficulty Levels . . . . .	61
VI. Latin Square, Group, and Number of Students Participating from Each Class . . . . .	64
VII. Two-Way Analysis of Variance Comparisons of Two Equivalent Tests at Three Difficulty Levels . . . . .	73
VIII. Summary of Raw and Adjusted Mean Transcription Error Scores for Four Latin Squares . . . . .	75
IX. Average Differences Between Overall Adjusted Mean Transcription Error Scores for Three Difficulty Level Comparisons . . . . .	76
X. Z Scores for Three Difficulty Level Comparisons . . . . .	77
XI. Relative Achievement of Groups Within Squares . . . . .	79
XII. Two-Way Analysis of Variance for Two Equivalent Tests and Three Difficulty Levels Using Grand Means . . . . .	79
XIII. Summary of Raw Scores for All Students on Six Study Tests . . . . .	116
XIV. Summary of Raw Mean Transcription Error Scores for Eight Classes . . . . .	120

Table	Page
XV. Computation of Z Scores for Three Difficulty Level Comparisons . . . . .	122

FIGURE

Figure	Page
1. Layout of Experimental Design . . . . .	67

## CHAPTER I

### THE RESEARCH PROBLEM

#### Introduction

Measuring student achievement is an integral part of the shorthand teaching process. As Hillestad (23, p. 1) indicates: "Reliable measurement of performance is an essential part of skill development." Most educators agree that such measurement should be based on established standards using testing instruments that accurately evaluate student learning.

Much of the evaluation in shorthand dictation classes is based on the student's ability to record shorthand notes from dictation at a given speed for three or five minutes and transcribe them with no more than 5 percent error. According to Hillestad (23, p. 1): "On the basis of such shorthand dictation tests, job placements are made, achievement standards are set, progress in class determined, and honors and awards are presented."

In addition, the results of many research studies in shorthand are based on achievement scores on shorthand dictation tests. If the tests given to students in such situations are not of verifiable difficulty or their comparability cannot be established, the findings of those studies are in question.

Because of the importance of such tests to teachers, students, employers, researchers, and others, valid and reliable testing instruments that are of known difficulty are necessary. Much has been written about the importance of being able to determine copy difficulty if shorthand dictation tests are to adequately and accurately measure student achievement.

Adams (1, p. 8) said that "each test at any specified speed should be consistent in difficulty to provide a reliable measure of student ability and growth." Likewise, Patrick and Hess (22, p. 248) indicated that "tests which are not of comparable difficulty do not portray accurately student's gains in shorthand writing speed." Pullis and Nickerson (59) further stated:

Shorthand teachers would agree that measurement in shorthand should be so designed as to indicate a change in the student's skill and not merely represent a divergence in the difficulty of the dictation material (p. 11).

And Hillestad (23, p. 1) agreed when she said: "Therefore, if achievement standards are to be meaningful, the testing materials used to measure students' skill in recording shorthand dictation should be of known difficulty."

In the past, syllabic intensity has been the primary measure of copy difficulty. However, a number of studies by researchers such as Wellman (74), Flood (15), Curtin (7), Turse (71), Uthe (72), Mellinger (43), Mickelsen (45), Pullis (53, 54), Henrie (20), and Nickerson (47) have shown it to be an inadequate index. Their findings were summarized by Adams (1, p. 8) when he stated: "Researchers over the years have reported that syllabic intensity is only one ingredient of a valid reliable shorthand test." And Pullis (56) cautioned teachers about its

influence when he said:

Do not assume that dictation takes of identical syllabic intensity are equated in difficulty. Changes in student performance may be reflecting changes in the difficulty of dictation materials rather than reflecting changes in student proficiency (p. 87).

Several other factors have been suggested as possible alternatives. One of the most commonly mentioned of these is vocabulary level. Typical of many comments, Pullis (58, p. 11) said: "Difficulty of shorthand material is more directly related to vocabulary frequency than to syllabic intensity."

The concept of vocabulary or word frequency assumes some method for identifying the words used more often than others and a subsequent ranking by order of use. One of the earliest lists of this type was developed by Timothy Bright in the 16th Century. Later lists were developed by Kaedig (30), Eldridge (11), Thorndike (67, 68), and Dewey (9) using a variety of sources and intended for various uses.

In 1926, Horn (25) developed the first specialized office vocabulary list which was later expanded in his work with Peterson (26).

The study by Silverthorn (63) completed in 1955 has been widely used in business research. From a sample of 2,039 items of correspondence drawn from 15 business categories in 41 states, Silverthorn identified a basic vocabulary of written business communication. This list represented over 300,000 running words and reinforced earlier findings that a large percentage of all correspondence is composed of a relatively small number of words.

In 1968, Perry (50) developed an updated list of words as part of his study to identify the most frequently occurring phrases in office correspondence. From a 317,306 corpus of words found in 2,061 pieces

of correspondence from a variety of types of businesses in every state but one, he ranked 12,109 words in order of use. From this, he determined the percent of total occurrence of various groups of words and found that the first 100 words represented over one-half, or 53.43 percent, of the words used in business correspondence. He also found that the first 500 words accounted for 71.93 percent, the first 1,000 accounted for 80.66 percent, the first 1,500 represented 85.58 percent and that more than 96 percent of the words used were within the first 5,000.

Perry (50) also found that the average typewriting stroke intensity of words was 5.8 including space and punctuation marks and that the average syllabic intensity was 1.63, which is much higher than the 1.4 in current use. He suggested that the higher syllabic intensity be adopted and that teachers "take the responsibility to see that students thoroughly master the 1,000 most frequently used words" (p. 109).

A similar study was completed by Mellinger (41) in 1970 based on a 295,271 word corpus. Mellinger also used stratified samples of correspondence drawn from a representative sample of 5,000 businesses, schools, and non-profit organizations throughout the nation. Upon comparison, his listing of 12,897 words was found to be essentially the same as that of Silverthorn (63) and Perry (50) except for minor variations in rank and inclusion of several words.

Both Hillestad (23) and Uthe (72) found that vocabulary level was an important factor in determining the difficulty of shorthand dictation materials. Wellman (74) and Flood (15) also concluded that vocabulary level or frequency of usage seemed to have an impact on copy difficulty, and Mellinger (43) suggested that a word frequency index based on

frequency of occurrence might be a better indicator of copy difficulty than syllabic intensity.

Reporting in the 1978 National Business Education Association Yearbook, Smith (65, p. 191) stated: "Vocabulary level appears to have a bearing on the difficulty of copy; however, additional work is needed to refine this factor in terms of copy difficulty."

West (76, p. 24) contended that stenographic tests should be a "good predictor of later life behavior," and that teachers should "match the content and conditions of school testing to the content and conditions of real-life uses of typewriting and stenographic skills" (p. 25). To accomplish this, he suggested, among other things, that there be "better equalizing of difficulty of copy materials via a standard word of 1.54 syllables plus a vocabulary index" (p. 32). Adams (1, p. 8) also suggested that difficulty of testing instruments be related to percentages of words as they occur in the "world of work."

It was the focus of this study, therefore, to determine whether copy difficulty could be measured and/or varied by using controls that closely simulated the vocabulary levels of real world conditions.

#### Need for the Study

Since the major objective of shorthand instruction is to help students develop the ability to record shorthand from dictation at a vocational level of skill, teachers must be able to evaluate the achievement of students both as they progress through the instructional process and at the end of a period of training. The accuracy with which such evaluations can be made depends to a large degree on the testing instrument used. If the test is too easy, student skill is overrated; if it is



too hard, students cannot pass the tests at speeds equal to their ability.

The usefulness of a measure of copy difficulty to teachers is indicated by Hillestad (23, p. 34) as follows: "If the difficulty of tests were measurable, equating tests would be possible, thus making standards in shorthand testing more meaningful than they seem to be at present."

Gallion (17) also stated:

Dictation materials for both practice and testing purposes should be based upon materials of comparable levels of difficulty. Using material of varying and unknown difficulty does not provide a sound basis for assessing the achievement of students (p. 44).

Further, Pullis (53) claims:

The inability to classify dictation material according to difficulty makes it impossible for educators to know whether measurement in shorthand has been so designed as to indicate a change in the student's skill or whether it merely represents a divergence in the difficulty of the dictation material (p. 1).

The lack of valid and reliable testing instruments also hampers researchers; and, unfortunately, the situation described by Haggblade (19) in 1965 has not changed today. In explaining why he developed and attempted to validate his own testing materials, he said: "It must be emphasized that no standard, valid tests for measuring shorthand achievement are available" (p. 30).

Bogges (4) also commented on the lack of a dependable measure of achievement by stating:

Since the value of the present study depended, in part, on the validity of using vocabulary level as an indicator of the difficulty of shorthand dictation material, it seemed to the present researcher at this point in the study that this type of study should not have been conducted yet because of the present lack of a reliable, consistent measure of the difficulty of shorthand dictation materials (p. 62-63).

Even though several researchers including Wellman (74), Elsen (13), Turse (71), Flood (15), Curtin (7), Hillestad (23), Farmer (14), Mellinger (42, 43), Petersen (51), Baggett (2), Uthe (72), Meyer (44), Boggess (4), Mickelsen (45), Henshall (21), Henrie (20), Pullis (53, 54), and Nickerson (47) have attempted in one way or another to identify factors that affect copy difficulty in shorthand dictation materials, the problem of determining difficulty levels has still not been resolved.

According to Pullis (55):

No one has yet been able to equate copy with such precision as to say with assurance, for example, that one set of letters marked for dictation at 80 words a minute is of the same difficulty as another set of letters marked for dictation at 90 words a minute (p. 156).

This dilemma is further reinforced by Nickerson (47, p. 94) who reports: "Even though several approaches have been used to determine factors relating to the ease or difficulty of shorthand dictation materials, there exists no conclusive measure of difficulty of dictation materials."

In the most recent National Business Education Association Yearbook, Smith (65, p. 190) said: "Much research is needed to determine factors that have a bearing on the difficulty of copy." He further said: "Research is needed to determine means of identifying the validity and reliability of test instruments for assessing terminal recording and transcribing achievement" (p. 191).

Finally, Pullis and Nickerson (59), reviewing a number of research efforts to identify factors contributing to copy difficulty, said:

In shorthand, we are unable at the present time to equate or predict with a high degree of accuracy the difficulty of typical dictation material as measured by the number of shorthand transcription errors committed (p. 13).

### Purpose of the Study

The purpose of this study was to identify criteria that could be used to develop copy of known difficulty for shorthand testing materials. This information would make it possible to consistently and accurately determine the difficulty of copy used for testing instruments and, hence, enable teachers to obtain a more adequate and meaningful measurement of student achievement. In addition, the results of research studies comparing achievement could be more accurately evaluated.

### Statement of the Problem

This study was undertaken to investigate whether word frequency could be used to determine copy difficulty for shorthand testing materials. Answers were sought to the following questions:

1. Is it possible to write shorthand testing materials that are of comparable difficulty by controlling the percentages of words used in various frequency categories while holding brief forms, brief form derivatives, syllabic intensity, number of different words, and number of actual words constant?
2. Can the difficulty of those tests be varied by changing the percentages of words in the different frequency categories by approximately 15 percent while holding brief forms, brief form derivatives, syllabic intensity, number of different words, and number of actual words constant?

## Hypotheses

To determine whether shorthand dictation tests could be developed that were similar in difficulty, the following null hypotheses were tested:

1. There is no statistically significant difference in average difficulty between two easy tests as measured by mean transcription error scores.
2. There is no statistically significant difference in average difficulty between two average tests as measured by mean transcription error scores.
3. There is no statistically significant difference in average difficulty between two hard tests as measured by mean transcription error scores.

To determine whether the difficulty of shorthand testing materials could be significantly changed by varying the percentages of the different categories of frequently used words, the following null hypotheses were tested:

4. There will be no statistically significant difference in student's mean transcription error scores on the easy difficulty level tests and the average difficulty level tests.
5. There will be no statistically significant difference in student's mean transcription error scores on the average difficulty level tests and the hard difficulty level tests.
6. There will be no statistically significant difference in student's mean transcription error scores on the easy

difficulty level tests and the hard difficulty level tests.

All hypotheses were tested with the probability of a Type I error equal to .01.

#### Delimitations

1. The shorthand system used in this study was limited to Gregg Shorthand, Diamond Jubilee series.
2. The study was limited to students enrolled in post-secondary institutions during Fall Semester, 1978, within the State of Oklahoma.
3. All percentages used for word frequency counts were based on the findings of the Perry (50) study.
4. Words used in constructing the tests for this study were taken from the various categories in the Perry (50) study.
5. Only shorthand related errors made on the transcript were considered. The scope of this study did not include an evaluation of student's shorthand outlines or nonshorthand errors.
6. No attempt was made to identify or classify students by IQ, age, sex, socio-economic, cultural, or ethnic background.
7. These materials were tested on students writing at least 80 but less than 100 words per minute. Differing effects on other speed levels, if any, were not considered.
8. Transcription time or rate was not a consideration in this study.

### Limitations

1. The effect of dictation presented on tape by an unfamiliar dictator is not known. However, students involved in the study had previously received practice dictation by tape from a variety of dictators. They had also had some experience with the test procedure and the investigator's voice on the preliminary test.
2. The effect of using intact classes rather than random assignment of students to treatments is not known. However, statistical methods were used to modify this effect.
3. The effect of holding brief forms, brief form derivatives, number of actual words, and number of different words constant on the tests used in this study is not known. Since all tests were treated alike with regard to these factors, however, their impact on the results should be similar.
4. The effect of limiting participants in this study to post-secondary students in structured classes writing 80 to 100 words per minute is not known. However, those participants represented a broad and diverse range of student characteristics and abilities from several schools with a variety of teachers.
5. The effect of using preliminary tests controlled only by syllabic intensity on the selection of participants is not known. However, all students were measured with the same tests and were, therefore, presumably of equal ability.

## Definition of Terms

In order to assist in the interpretation of this report, the following terms are defined as they were used in this study.

Brief forms--Abbreviated outlines for certain words that are used in the Gregg system to facilitate more rapid writing.

Brief form derivatives--Words which include brief form outlines.

Difficulty level--The comparative degree of difficulty of a shorthand test item on the scale of easy, average, and/or hard.

Easy--A three-minute dictation item in which the percentages of words identified by Perry (50) have been adjusted to contain approximately 15 percent more words in the high frequency groups shifted from the low frequency groups.

Average--A three-minute dictation item in which the percentages of various categories of words conform to those identified by the Perry study.

Hard--A three-minute dictation item in which the percentages of words identified by Perry have been adjusted to contain approximately 15 percent more words in the low frequency groups shifted from the high frequency groups.

Number of actual words--The total number of individual words found in each of the six tests developed for this study.

Number of different words--The total number of words which are used at least once within each of the individual tests developed for this study.

Shorthand transcription errors--Errors in the shorthand transcript which include addition, omission, or transposition of words which vary from that dictated.

Nonshorthand transcription errors--Errors in the shorthand transcript not related to the translation of shorthand outlines into typewritten copy such as spelling, grammar, punctuation, or typographical errors.

Perry word list--A listing of 12,109 different words by their frequency of occurrence based on an analysis of 2,061 different business letters incorporating 317,306 word occurrences.

Syllabic intensity--The average number of syllables per word in a dictation item as determined by dividing the total number of syllables by the total number of actual words.

Test take or dictation take--A three-minute item of dictation composed of unfamiliar, new material.

Transcript--The typewritten translation of shorthand symbols.

#### Assumptions

1. Perry's word list is a valid indicator of the frequency of words used in business correspondence.
2. The percentages indicated for the various categories of words in the Perry list are representative of typical business correspondence.
3. A three-minute dictation test is an adequate measure of shorthand skill.
4. The tests used for the preliminary testing of students are as adequate as any presently available for this purpose.



## Summary

Despite a great deal of research in the past, there is still an acute need to identify a means for measuring and/or controlling copy difficulty for shorthand testing materials. Because there is some evidence in the literature to indicate that vocabulary level may have a bearing on this problem, this study sought to determine whether tests using similar percentages of words in various frequency categories would be of comparable difficulty and whether copy difficulty could be significantly changed by varying the percentages of words in various frequency categories by approximately 15 percent while holding brief forms, brief form derivatives, syllabic intensity, number of different words, and number of actual words constant.

The remainder of this study is organized in the following manner: Chapter II contains a discussion of previous studies relating to copy difficulty. Chapter III is a summary of the procedures used in conducting this study, and Chapter IV reports the findings obtained. Finally, Chapter V contains a summary of the study together with conclusions and recommendations.

## CHAPTER II

### RELATED LITERATURE

#### Introduction

The need for tests of known difficulty to measure shorthand achievement has been evident for many years. Since the late 1920's, many approaches have been taken and a variety of factors have been tested as researchers have attempted to find a method to measure the difficulty of copy used for testing purposes. Some of these studies have dealt with characteristics of the dictation material, others with factors of the shorthand system. However, as yet, no conclusive evidence has been found which identifies the factor or factors that can be used to measure and/or control copy difficulty for shorthand testing materials.

There is some evidence that word frequency as identified by several word lists has an impact on the difficulty of copy. Several writers and researchers have indicated that the achievement of students is affected by their ability to write the outlines for high frequency words quickly and accurately, and previous studies have found some relationship between word frequency and copy difficulty.

Since the reports of several other researchers during the past few years have reviewed the literature grouped by factors and variables studied, and to provide some historical perspective to this problem, this chapter will present a chronological review of a number of studies

dealing with various attempts to test factors that could be used to determine copy difficulty.

### Studies Dealing With Copy Difficulty

#### Schrampfer, 1927

One of the earliest formal studies done on copy difficulty was conducted by Schrampfer (62) in 1927. Three letters with shorthand stroke counts of 2.32, 3.53, and 4.59 respectively were given to first, second, third, and fourth semester high school students in Iowa. Schrampfer found that the third letter was the most difficult for students at all levels to write. Although word frequency was not considered a factor contributing to difficulty in the conclusions of the study, the following observation was made:

The first letter, Test One, consists of very common words, many of which have probably been made automatic by the student. Test Two contains words of greater difficulty, and Test Three has many words that are little used and little drilled upon in the classroom (p. 123).

Even though findings of this study indicated that shorthand stroke count was a more reliable measure of difficulty than one based on words or word units, Schrampfer (62, p. 126) recognized that "mental difficulty with unfamiliar words" hampered the usefulness of her method and suggested that some adjustment be made for uncommon words.

#### Leslie, 1931

In 1931, Leslie (36), following a recommendation by Gregg, proposed syllabic intensity as a measure of copy difficulty and advocated a standard word of 1.4. This figure was arrived at from an analysis of New York

State Regents Examinations, shorthand speed contest materials, dictation test materials of the Gregg Publishing Company, Congressional Record material, and word studies available at that time. It is interesting to note that Leslie (38, p. 282) later recognized the impact of vocabulary on copy difficulty when he stated "the difficulty of dictation material whether in terms of dictation speed or difficulty of transcription varies in accordance with the spread or range of the vocabulary." However, he chose syllabic intensity over a vocabulary measure because it was simple to compute and because, as he stated, "empirically we know that the syllabic intensity gives us immediately a good estimate of the difficulty of ordinary English running material because it also gives us a good estimate of the spread of the vocabulary (p. 282).\* According to Leslie (36), a high syllabic intensity in dictation material composed of ordinary business dictation indicated the presence of a greater number of low-frequency, and supposedly harder, words.

Leslie (37) also reported that a number of other attempts had been made to determine and control copy difficulty for shorthand dictation material. Such factors as typing stroke count, shorthand character count, sound count, artificial restriction of vocabulary, vocabulary spread index, vocabulary analysis, sentence length, syllabic intensity, and the standard word had been investigated.

While admitting that the difficulty of the materials was not the only variable influencing difficulty in recording, Leslie (35) advocated

---

\*This assumption was later challenged by Hillestad (23) who pointed out that dispersion (vocabulary spread) cannot be shown with a measure of central tendency (syllabic intensity). She further showed that the relationship between syllabic intensity and vocabulary spread is not linear.

further research to discover a way to measure the material simply and easily.

Wellman, 1937

Another early study was conducted by Wellman (74, p. 2) in 1937 in an attempt to find "some formula for evaluating dictation material." Dealing with the factors of syllabic intensity, vocabulary frequency, shorthand stroke intensity, and sentence length, she composed ten 150-word letters. One factor in each test was varied and the other factors were held constant. Seven tests included in a group called Battery A and three tests included in a group called Battery B were dictated to advanced high school and college stenography students, and mean number of transcription errors on each of the ten letters were compared for significant differences.

Results indicated that there was no significant difference when the length of sentences was doubled, shorthand stroke content was increased by 20 percent, syllabic content was increased by 28 percent, an increase of 17 percent in syllabic intensity was combined with an increase of 20 percent in stroke intensity, or 12 percent of the vocabulary was shifted from the 1,000 most frequently used words to higher levels within the 3,000 most frequently used words. She did, however, find a significant difference between two letters in which all elements were alike except that 30 percent of the vocabulary was shifted from the 1,000 to the 2,000-5,000 most frequently used ranks of the Horn (26) list. She also found an increase in difficulty when spelling demons were incorporated. From these findings, Wellman concluded that vocabulary level appeared to be a better measure of difficulty than syllabic intensity, number

of words, number of strokes, and number of occurrences of ranked word forms. She further suggested that "elusive qualitative elements, intrinsic to the thought and language content of the dictation materials are more likely determinants of stenographic difficulty than are the physical components that can be measured and objectively scaled" (p. 87).

Wellman recognized the limitations imposed by the lack of statistical and quantitative measures for handling the data for her study, and had she had access to more modern analytical procedures would not have had to limit the vocabulary and other factors as she did. Uthe (72) recommended that the results of this study be generalized cautiously since the two batteries of tests were given to different groups and no randomization was indicated.

#### Elsen, 1946

In contradiction to Wellman's (74) findings, syllabic intensity was determined to be a better predictor of difficulty than the factors of vocabulary spread and shorthand character count by Elsen (13) in 1946. Two series of six items consisting of five letters and half of a Congressional speech in each series were dictated at 80 words per minute to 100 high school seniors. The syllabic intensity of the letters ranged from 1.3 to 1.7 and the syllabic intensity of the Congressional speech was 1.9. (The syllabic intensity of 1.8 was omitted.)

Syllabic intensity and number of errors made on each item were then plotted on two-line graphs for each series. Based on visual inspection of the graphs, Elsen claimed that syllabic intensity had the greatest validity of the three factors studied for predicting the difficulty of shorthand dictation material.

Several problems inherent in this study are pointed out by Hillestad (23) and Uthe (72). First, the units for syllabic intensity and the units for errors were shown on separate scales but presented in the same graph. Second, no statistical analysis other than visual inspection of the two lines was reported; and third, no statistical tests of significance were carried out on the difference in errors made on the two sets of dictation. Uthe also commented that "the degree of relationship between syllabic intensity and errors was not considered by Elsen, although validity of an instrument is the degree of relationship between that instrument (the predictor, syllabic intensity) and the criterion (errors)" (p. 16). It should also be noted that random selection of the letters was not indicated, and students had apparently had little experience with material such as the Congressional speech.

#### Turse, 1944-1948

The shorthand stroke was suggested by Turse (70) in 1944 as a means of determining shorthand difficulty. He proposed that a combination of shorthand word count, actual word count, and shorthand stroke be developed to increase the predictive power of syllabic intensity. If this could not be done, he recommended that one or a combination of those that came closest be used.

Later, in 1948, Turse (71) conducted an informal study using three letters of 150 words each dictated to 144 stenographic job applicants in New York City at speeds of 60, 80, and 100 words a minute. Word length, syllabic count, and shorthand stroke count were the factors used for indicating difficulty. Analysis of the results showed that none of the correlation coefficients was greater than .28, indicating that none

of the three variables were valid measures of difficulty. Turse observed that the number and kinds of shorthand strokes should be considered in determining difficulty and that the contextual setting of a word appeared to make a difference to the writer.

#### Flood, 1953

That frequency of use of a word may be a better indicator of the difficulty of dictation material than syllabic intensity was suggested by Flood (15) in 1953. After comparing the learning loads of two systems of shorthand, she concluded that "the many inconsistencies in the application of the principles (of Gregg) increase the memory burden unnecessarily and may tend to create hesitation and confusion in the writing of new words" (p. 169). She questioned the practice of using syllabic intensity as the sole determinant of copy difficulty and believed that long words that were frequently used might be easier to learn than short ones that were unfamiliar.

#### Wessman, 1956

Three years later, Wessman's (75) study indicated that there was, in fact, a relationship between word frequency and difficulty as measured by errors made in shorthand notes. After dictating two items with the same syllabic intensity to 72 students who had been randomly selected from four shorthand classes in three schools, he found successive increases in errors as the vocabulary frequency by group decreased. When considering errors by frequency group, he found that 73 percent of the errors were made in the 28 percent of the words that had the lowest frequency.

In addition, even though the two letters had the same syllabic



intensity, he found a higher percentage of errors were made on the second letter than the first. Further analysis revealed a statistically significant difference between the two letters in the distribution of the words by frequency group with the second letter having more words in the less frequently used categories. From this information, the conclusion was drawn that students have more trouble writing the less common words.

Curtin, 1958

In an attempt to determine other factors having an effect on the difficulty of dictation materials, the Cloze procedure used in the language arts field as a measure of readability of both oral and written communications was tested by Curtin (7). Following a procedure developed by Taylor (66), every fifth word was omitted in three letters, and students were instructed to fill in the blanks by guessing. The more correct responses, the higher the Cloze score and the easier the material was thought to be.

She also dictated a sample of 41 letters randomly selected from Dictation for Mailable Transcripts to nine classes of second-year shorthand students to obtain the number of shorthand errors and derive predictor scores for each letter. Curtin assumed that if the Cloze score were able to predict difficulty, there would be an inverse relationship between the Cloze score and shorthand errors. Using a vocabulary level index developed by Hillestad (23), together with syllabic intensity, correlations were made between and among the shorthand error scores and the predictor scores for the Cloze procedure. No significant relationship was found between the Cloze score and the number of errors made on the shorthand notes. A nonsignificant coefficient of correlation of .128 was also

obtained for the comparison between syllabic intensity and the number of shorthand errors indicating that syllabic intensity would not predict difficulty. The correlation coefficient between the vocabulary level index and errors made on the shorthand notes was .501, and even though this figure was not significant, Curtin (7) concluded that vocabulary level appeared to be a better predictor of difficulty than either the Cloze score or syllabic intensity.

#### Danielson, 1959

Another study designed to determine the relationship between shorthand vocabulary competency and achievement in shorthand dictation was completed by Danielson (8) in 1959.

Choosing 50 words from each 1,000-word level in the Silverthorn (63) list, Danielson constructed six word-list tests of 250 words each, which were given to 120 university students over a period of two semesters. Beginning in September, one of these tests was given every six weeks. Students were instructed to write the dictated words in shorthand and transcribe them on the typewriter. The number of words transcribed correctly became the vocabulary index.

For the second part of the test, Danielson constructed 30 sets of business letters with a syllabic intensity of 1.5. One set of these letters was dictated each week at speeds ranging from 60 to 130 words per minute. Students were required to pass three takes to establish their dictation rate.

Analysis of the results revealed that shorthand vocabulary competency was significantly related to achievement in shorthand dictation indicating that as a student's shorthand vocabulary competency increased

the rate of taking dictation also increased. Uthe (72), however, questioned the equality of the tests used in this study since syllabic intensity had previously been shown to be an inadequate measure of difficulty; and, therefore, all tests could not be considered to be the same.

#### Crandall, 1960

Another approach to the impact of word frequency on copy difficulty was taken by Crandall (6) in an informal study. His assumption was that the more frequently used words would have become easier to write and transcribe through repeated practice, and the less frequently used ones would be harder since they had been written less.

Using Elsen's (13) results showing syllabic intensity and transcription errors per student for six tests together with the Silverthorn (63) list, Crandall computed a composite word frequency index for each letter expecting that there would be a positive correlation between his frequency index and the shorthand errors reported by Elsen. The results did not bear this expectation out, however. Speculating that a few words might be distorting the findings, a second analysis was made using 100-word blocks to develop the vocabulary index; and this time a positive relationship between word frequency and transcription errors was obtained. Since the first block of 100 words accounted for over 50 percent of words used in business communication, with the remaining 50 percent spread over several 100-word blocks, Crandall concluded that transcription errors were more directly related to the half of business communication composed of infrequently used words than frequently used ones, pointing out that the ability to write the less frequently used words was, therefore, of major importance to the learner.

Hillestad, 1960

With the advent of the computer, Hillestad (23) sought to develop a multiple regression equation that would identify characteristics of the shorthand system and/or dictation materials that affected ease or difficulty of writing as measured by errors in students' shorthand notes.

To do this, she developed 100 letters that were especially written to include measures of several variables, some of which were related to the dictation materials and some of which were related to shorthand principles. Each letter was 160 words long and had been revised from letters collected from a variety of businesses to test the following factors: syllabic intensity; vocabulary level index; brief forms; brief form derivatives; blends; vowel sounds including o, oo, and diphthongs; terminal t after k or s; plurals and past tenses, suffixes, and prefixes. These letters compared in distribution of variables with a sample of letters from Dictation for Mailable Transcripts.

The letters were dictated by the teachers over a three-month period at a speed students could comfortably write to eight high school fourth-semester classes who were writing Gregg Simplified. All letters were randomly placed in blocks, and the blocks were randomly assigned to classes. To eliminate the effects of learning during the testing period, all students took all letters. Five papers were randomly chosen from each of the eight classes providing a sample of 40 papers checked for each letter. Only one error per word was counted; and errors were not counted for placement on the line, size, proportion, or reversal of letters unless it changed the letter.

A multiple correlation coefficient of .948 was obtained for the

relationship between the 16 variables and the error scores. Each of the 16 variables was then tested for predictive ability, and six variables remained significant. These six variables were then tested for correlation with error scores, and a multi-correlation coefficient of .943 resulted indicating that the six variables could predict difficulty almost as well as all sixteen.

Further analysis revealed that the two variables called "syllabic intensity" and "vocabulary level index" accounted for 73.36 percent of the criterion variance, and only 15.56 percent was due to variance caused by the other four of the six variables.

Substituting the variable entitled "words beyond 1,500 on the Silverthorn list" for "vocabulary level index," Hillestad (23) obtained a coefficient of .78 for the relationship between the two remaining variables and error scores leading to the conclusion that even though there was a significant difference at the .01 level between the predictive ability of the equation with 16 variables and the one with only two, the last one adequately predicted difficulty and was much simpler to compute.

An analysis of the errors made by students revealed that the variable "words beyond the first 1,500 on the Silverthorn list" contributed twice as much to the variability in number of errors as did the variable "syllabic intensity." She also found that students had lower error scores on brief forms than words constructed according to principles of shorthand and that as words became longer, the error rate tended to increase. Inconsistently applied shorthand principles and word endings also caused problems in writing for students.

It appeared from the findings that the less frequently the words were used, the greater the percentage of error on them since less than

15 percent of the errors were made on the first 100 words of the Silverthorn (63) list, and five times that many were made in the category consisting of the words ranked from 101-300. In addition, the percentage of error continued to increase from 3 to 6 percent with each higher level to a peak of 42 percent error on words ranking beyond 5,050 on the Silverthorn list.

Hillestad (23) recommended broader vocabulary coverage with more practice on the less frequently used words. In addition, she stated:

In selection of dictation materials, both for practice and for testing, more attention should be paid to the number of words beyond the first 1,500 on the Silverthorn list contained in the dictation material. The number of these words in a piece of dictation seemed to be a better indicator of difficulty than did the number of syllables the words contained (p. 118).

Hillestad also advocated the automatization of a greater number of the first 500 words on the Silverthorn list and more work with brief form derivatives. She suggested that help be given to students in distinguishing between the oo and o sounds, that teaching materials be organized so that the easier principles would be covered first, that those materials be arranged according to frequency of use of words, and that principles of shorthand also be coordinated with the frequency of use of words.

As has been pointed out by other researchers (72) (47), the fact that no controls were applied to the speed of dictation may have had some impact on the results of Hillestad's study. The letters may not necessarily have been representative of business correspondence, and only shorthand notes were considered. In addition, in the selection of the sample of five papers for each letter from each school, some students were selected for more than one letter, causing an undetermined amount of dependence in the error scores for each letter.

Farmer, 1961

Approximately a year after the completion of Hillestad's (23) study, Farmer (14) attempted to validate the prediction equation using transcription scores as the criterion. From Hillestad's 100 letters, Farmer chose 83 that had approximately the same syllabic intensity as published dictation materials. Using syllabic intensity and words beyond the first 1,500 on the Silverthorn (63) list as criteria, she grouped the 83 letters into high, medium, and low categories. She then chose two letters from each category making a total of six letters which were dictated at 60 words per minute to 96 students in six classes of second-year Pitman shorthand.

An analysis of variance was used to test for significance of mean transcription errors in classes, among difficulty levels, and for interaction. Farmer found that the letters classified as difficult were significantly more difficult than either the medium or easy classifications. The mean number of transcription errors was 73.69, 44.19, and 47.06 for the three respective classifications. Because the difference in difficulty between medium and easy was not significant, however, the conclusion was drawn that the formula had not been successful in determining the three levels of difficulty.

Farmer urged further study to validate Hillestad's formula, pointing out that the conclusions of this study may reflect differences caused by the fact that Hillestad developed the formula from materials written in Gregg Simplified and her study used Pitman. Farmer also pointed out that one class in which the better students were given the easy tests and the poorer students were given the medium tests may have contributed to the lack of a significant difference between the easy and medium letters.

Peterson, 1964

In 1964, Peterson (51) also attempted to validate Hillestad's (23) formula. Eight original one-minute letters with a syllabic intensity of 1.4 and emphasizing different levels of 400-word groups from Silverthorn's (63) list from 100 to beyond 5,000 were dictated at 50 words per minute to a random sample of 100 first-year and 100 second-year Gregg shorthand writers. Using transcription scores as the criterion variable, Peterson found wide error variances within the three levels and concluded that although there was some evidence of relationship between difficulty level and percent of error, the results of his study did not support the predictive ability of Hillestad's formula. Uthe (72) questioned the technique of using an irregular progression from one vocabulary level to another, however, and pointed out that this study must be viewed cautiously because of the brevity of the letters and the lack of information about the similarity of the classes.

Baggett, 1964

Baggett (2) also attempted to validate Hillestad's (23) formula that same year and also used transcription errors as the criterion. Six letters were selected from those used by Hillestad in the predicted order of increasing difficulty from 1 through 6. These letters were then dictated at 80 words per minute to 600 students writing Gregg Simplified in seven fourth-semester high school classes. From these papers, Baggett randomly chose 50 transcripts and computed the mean transcription errors which ranged from 2.31 to 8.65. Even though there were differences in the scores, those differences were not in the order expected when the



Spearman Rank Correlation Coefficient technique was applied. The final results ranked the six letters in an order of 2, 3, 4, 1, 6, 5 instead of the 1 through 6 Hillestad's (23) formula would indicate. Baggett (2) offered several explanations that may account for some of the discrepancies in letter order he found. Among them were first, the possible effect of class differences since the sequence of difficulty found among the letters was not the same for all classes; second, sentence length, since one letter had only three sentences in 160 words; and third, some confusion caused by incorrect grammatical construction. He also suggested that further study be done with other factors such as sentence length, series and order of words, and vocabulary level to determine their effect on copy difficulty.

#### Mellinger, 1964

Challenging syllabic intensity, Mellinger (43) proposed a word frequency index to determine copy difficulty in 1964. Citing a study in which he analyzed Silverthorn's (63) word frequency list to determine whether the 1.4 syllabic intensity in common use was reliable, Mellinger (42) reported his finding that the average syllabic intensity of all 300,000 words was actually 1.56 and rose to 2.2 when the first 200 words were eliminated.

He further pointed out that even if a 1.5 syllabic intensity were to be adopted as a criterion, the difficulty of copy would still not be equated since students could not write low frequency words of the same syllabic intensity as fast as those with high frequency. This is illustrated in the ease with which high frequency monosyllabic words such as "the," "your," and "here" are written contrasted to low frequency

monosyllabic words such as "axe," "brooch," "lump," and "realm." The syllabic intensity measure assumes all monosyllabic words are equal in difficulty whereas a word frequency measure recognizes their differences to the student.

Since the first 200 words account for 60 percent of the 300,000 word occurrences studied by Silverthorn (63), Mellinger (43) proposed a formula in which letters having 60 to 69 percent of their words from that category, be called "average." Letters with less than 60 percent of the words from the first 200 would be classified "difficult," and letters with 70 percent or more of the first 200 words would be termed "easy."

Like Crandall (6), Hillestad (23), and others, Mellinger also advocated more emphasis on infrequently occurring words, which he identified as those beyond 200 of the Silverthorn list.

#### Uthe, 1966

Because of the failure of previous studies to validate Hillestad's (23) formula when carried through to the transcription phase, Uthe (72, p. 1) attempted to develop a new formula in 1966 that would "consistently predict the difficulty level of dictation material used in learning and testing situations." Since the Gregg system had undergone a revision, she also sought to identify those principles of Diamond Jubilee that caused most difficulty to students in writing. Using Gregg Diamond Jubilee, a different statistical design than Hillestad, and 35 variables including Hillestad's 16 as well as others used by previous researchers, Uthe arrived at a multiple regression formula for predicting copy difficulty.

The variables used in this study were of two types: those

characteristics inherent in the words themselves or in the character of the dictation material, and those related to the shorthand system as shown below:

- Syllabic intensity
- Vocabulary level
  - Brief forms in the 1-100 vocabulary level range
  - Constructed words in the 1-100 vocabulary level range
  - Words in the 1-100 vocabulary level range
  - Words in the 1-500 vocabulary level range
  - Words in the 501-1,500 vocabulary level range
  - Words beyond the 1,500 vocabulary level range
- One syllable words
- Two syllable words
- Three syllable words
- Four- to six-syllable words
- Punctuation marks
- Sentence length
- Typing stroke intensity
- Shorthand stroke intensity in brief forms
- Shorthand stroke intensity in brief form derivatives
- Shorthand stroke intensity in constructed words
- Shorthand stroke intensity
  - Brief forms
  - Brief form derivatives
- Blends
  - oo hook
  - o hook
- Plurals
- Blend-past tense combinations
- Past tense (t or d only)
- Past tenses (all)
- Disjoined prefixes
- Joined prefixes
- Diphthongs
- All beginnings
- All endings (Uthe, p. 116)

Following a stepwise regression procedure, a regression analysis program determined those characteristics that were significantly related to the degree of accuracy in shorthand notes. Appropriate weights were assigned to each variable, and at each step in the analysis the least significant variable was dropped and a new equation was begun.

All variables remaining at the .95 level of probability, the partial regression coefficient and F value of each retained predictor, and the

multiple correlation coefficient for this linear combination of predictors with the error score were included in the final equation.

Those variables that remained at the .95 level of probability and those with significant F values were then used to develop the multiple correlation coefficient of the variables to the error scores, the beta of the regression line, and the predicted number of errors for each of Hillestad's (23) 100 letters. The three significant variables included brief forms, words beyond the first 1,500 of the Silverthorn (63) list, and word endings. The correlation between them and word error scores was .76 and the multiple coefficient of determination was .58. Using standard deviations above and below the predicted word errors, the categories of "easy," "average," and "difficult" were established.

The 100 letters previously developed by Hillestad were dictated on records at 80 words per minute and given in four-letter blocks to 25 groups of fourth-semester high school students in the St. Paul area according to a randomized block design. All students also received three "common" letters especially constructed for this study as a control. After discarding those papers on which more than five consecutive words were omitted, a sample of three papers was selected for each of the 100 Hillestad letters.

An error analysis using vocabulary level, length of word, and selected shorthand principles was done for both Gregg Diamond Jubilee and Gregg Simplified. These findings were also compared with results from Hillestad's study. The analysis revealed that students made more than twice as many errors on Diamond Jubilee than Simplified and that the percentage of error on brief forms tripled even though the number of brief forms decreased in the Diamond Jubilee revision.

To validate the equation, Uthe (72) ranked the 100 letters according to predicted error scores and chose six letters, four of which were classified average and two of which were classified difficult. No letters in the easy classification were used. Students in four classes in one school taught by the same teacher had previously been assigned by ability at registration. Class A students were above average in ability, Class B was average, Class C students were below average in ability, and Class D was average. Classes A, B, and C received the letters at 60 words per minute and Class D was given the letters at 80 words per minute.

When mean shorthand outline error scores for Classes A, B, and C were computed, the six letters all appeared in the predicted difficulty category. The mean shorthand outline error scores for Class D, however, was generally one level higher than expected. From these data, Uthe (72, p. 118) developed the scale for classifying dictation materials shown below:

<u>Errors</u>	<u>Difficulty Level</u>
0 - 7.11	Extremely Easy
7.12 - 15.49	Very Easy
15.50 - 23.87	Easy
23.88 - 32.25	Average (low)
32.26 - 40.63	Average (high)
40.64 - 49.01	Difficult
49.02 - 57.39	Very Difficult
57.40 +	Extremely Difficult

Uthe claimed that the scale shown above would determine the difficulty level for students of average ability writing at 80 words per minute and suggested adjustments of one standard deviation above or below to accommodate differences in ability or speed.

Both Hillestad (23) and Uthe found that vocabulary level (or words

beyond the first 1,500 of the Silverthorn (63) list) had a significant impact on copy difficulty. As Nickerson (47) pointed out, this importance is further strengthened by the variable "brief forms" in Uthe's (72) study since most brief forms are found in the first 500 words of Silverthorn's list and actually constitute another measure of vocabulary.

Meyer, 1967

The following year, Meyer (44) attempted to validate Uthe's (72) formula for determining the difficulty of dictation material. This study, by carrying the process through to transcription, also attempted to determine the extent to which shorthand errors result in transcription errors.

Twelve letters were selected from the 100 developed and used by Hillestad (23) and also later used by Uthe. There were four letters in each of the three categories of "easy," "average," and "difficult" as determined by Uthe's formula. The letters were recorded on tape and given to 95 fourth-semester high school shorthand students in eight different schools.

Using an analysis of variance procedure, Meyer found no significant difference in the groups from the eight schools. There were, however, significant differences in the errors on the 12 letters and in the three levels of difficulty. Significant differences were also found within each level of difficulty as well as some interaction of letters within groups.

The Scheffe' test was used to compare the mean transcription error scores and resulted in identification of "extremely easy" and "extremely hard" categories, but any other definite distinctions between those

levels could not be made. Meyer (44), therefore, concluded that her effort to validate Uthe's (72) formula using transcription error scores was inconclusive.

In an analysis of shorthand and transcription errors, Meyer identified several other factors that may contribute to the difficulty of dictation material. These were awkward wording of sentences, extremely high syllabic intensity, possible curvilinear relationship of syllabic intensity and difficulty, and subject matter and meaning of interest to the student.

Meyer also recommended further study of factors contributing to the difficulty of dictation material using published classroom materials rather than specially written letters to cover specific theory principles. Further study was also suggested to analyze dictation material for readability, sentence structure, contextual clarity, and the effect of high and low extremes of syllabic intensity.

#### Boggess, 1970

Finding existing formulas inconclusive and desiring to use the best predictor available, Boggess (4) conducted a study in 1970 in which she attempted to compare the achievement of students given practice dictation of varying levels of difficulty as measured by vocabulary level. Difficulty of materials in this case was based on a 6 percent increase and a 6 percent decrease in percentage of words above and below the first 1,995 of the Silverthorn (63) list. Since Silverthorn had determined that approximately 90 percent of the words in all business communications fell within the first 1,995 words on that list, the decision was made to consider material with 10 percent of the words above 1,995 as average. This

was further expanded to include 8 to 10 percent of the words above 1,995 in the average category. The easy category had 0 to 2 percent of the words above the first 1,995, and the difficult had 16 to 18 percent above. The 429 beginning shorthand students who completed the study were separated into four groups and given four different kinds of dictation. One group was given easy, one average, one difficult, and one a progression of easy to difficult. Fifty students were then chosen from each group for a total of 200 in the sample. Scores on three pretests--one easy, one average, and one difficult--indicated no significant difference in ability of any of the groups at the .01 level. Students received the special dictation during the last 12 weeks of the second semester. This was followed by a posttest similar to the pretest.

The posttests identified two levels of difficulty, but there were no significant differences at the .01 level among any of the groups from the pretest to the posttest. The conclusion was reached that the difficulty level of the practice material had little impact on the ability to take and transcribe dictation of all levels of difficulty.

Bogges (4) further concluded that the pretests were not sufficiently different from each other to warrant calling them "easy," "average," and "difficult" even though they had been determined by the same measures of difficulty as the material in the experiment. She cited two possible explanations for this situation. First, the measure of difficulty chosen (a measure of word frequency) may not be reliable for measuring difficulty of shorthand dictation materials; and second, the percentage of low frequency words used for each level was not sufficiently discriminating.

Bogges questioned the results of her study and stated that the



measure chosen to determine difficulty of shorthand dictation materials was not consistent and, therefore, not reliable.

Mickelsen, 1970

Another study in 1970 by Mickelson (45) determined that the relationship between vocabulary level and transcription errors tended to be inverse and that the error rate increased as vocabulary level decreased in frequency.

To arrive at these conclusions, Mickelsen constructed three 3-minute dictation tests with high frequency word indexes of 100, 70, and 40 based on the first 500 words of the Perry (50) word list. This meant that all words in the first letter were within the first 500. Letter 2 had 70 percent from the first 500, and the rest stratified by 500-word blocks according to the percentages in the Perry study. Letter 3 had 40 percent from the first 500 words with the remaining 60 percent stratified according to the percentages in the Perry list. Syllabic intensity in all three letters was held constant at 1.43.

The three letters were recorded on tape at 80 words per minute and given in randomized order over a period of two weeks to 117 fourth-semester high school students in three groups. All transcripts were hand scored counting only errors in transcription, and an analysis of variance and Dunn's "c" test were used to analyze the raw scores. Results indicated highly significant differences among raw mean transcription error scores for all three tests. In addition, every mean transcription error score was significantly different from every other mean transcription error score at the .01 level.

An error analysis revealed a general increase in errors as word

frequency decreased and that the greatest number of errors was found in two and three syllable words. The longest words did not result in the greatest percentage of error, and students tended to add a greater number of words as the difficulty increased.

Mickelsen (45) concluded that the use of high frequency words was very successful in establishing three distinct levels of difficulty in shorthand dictation material and that familiarity of vocabulary was a very significant factor in measuring the difficulty of shorthand dictation materials. Since transcription errors were directly related to vocabulary level, Mickelsen suggested that vocabulary level could be used as a single factor in determining the difficulty of dictation material for dictation purposes. He also concluded that syllabic intensity was an inadequate measure of difficulty and that 1.43 syllabic intensity did not allow a natural flow of wording similar to that used in normal business correspondence. Other factors seeming to impact difficulty were familiar words in unfamiliar settings and the change in listenability, coherence, and context as the indices of high frequency words decreased.

#### Henshall, 1970

Also in 1970, Henshall (21) sought to determine whether any one of a combination of four adult readability formulas could predict difficulty of shorthand dictation materials more accurately or more easily than the formula developed by Uthe (72). She also attempted to learn whether readability formulas would offer a more accurate internal discrimination of the levels of material difficulty than the shorthand formula.

Using the Uthe formula, Henshall chose 15 of Hillestad's (23) 100 letters covering the five difficulty levels identified by Uthe. These

letters were taped at 80 words per minute and given as five tests to 12 sections of beginning, intermediate, and advanced shorthand students at the college level. Errors on 71 transcripts of Gregg Diamond Jubilee writers were used in simple, multiple, and partial correlations of the difficulty levels predicted by the Dale-Chall, Flesch Reading Ease, Gunning-Fog Index, and the Farr-Jenkins-Patterson readability formulas with the Uthe (72) formula. While no one readability measure produced a simple correlation coefficient to measure difficulty as well as or better than Uthe, all combinations of the formulas were significantly predictive and a combination of the Gunning-Fog and Farr-Jenkins-Patterson formulas produced an  $r$  of .81 compared to an  $r$  of .84 for Uthe's prediction formula. Even though these correlation figures are very close, however, Uthe's formula was recognized as a better predictor when considering order of difficulty rather than specific error scores.

To evaluate ease of scoring, Henshall (21) obtained measures of the time required by 40 prospective shorthand teachers in methods classes and members of an undergraduate professional organization to apply the five formulas to each of three test letters. Ease of application was determined by the time required to compute the score, and an analysis of variance was used to determine significance of variances. Conclusions indicated that the readability formulas produced significantly superior results in terms of accuracy and that the time required was less for the readability formulas in every case.

Henrie (20, p. 34) reports that in personal correspondence, Henshall maintained that "one of the biggest drawbacks of the Uthe formula is the time required to figure the difficulty of a piece of material." Henshall

also claimed that the "difficulty of materials is a function of language and not so much of the system as had been previously claimed."

Henrie, 1971

Because of the different results obtained by using the various methods advocated for determining the difficulty of shorthand dictation material and the inability of subsequent researchers to validate any of them, Henrie (20) set out in 1971 to analyze and compare the four most prominent formulas which he identified as the syllabic intensity prediction formula used by Zoubek (78) and the difficulty prediction formulas developed by Hillestad (23), Mellinger (43), and Uthe (72) in an attempt to determine which one was most valid and reliable.

To accomplish this, he obtained a mean word error score which he used as a criterion by administering 20 two-minute taped letters ranging in speed from 70 to 100 words per minute to 13 fourth-semester shorthand classes. From these he chose 100 transcripts. Difficulty level scores were obtained by applying each of the four formulas to each of the 20 letters and then difficulty ratings ranging from very easy, easy, average, difficult, to very difficult were determined.

Comparisons were then made between ratings of each of the four formulas and the word error score, against each other prediction formula, and against each prediction formula itself. Analysis of variance, the Pearson product moment correlation, the split-half reliability formula, and a t-test for testing the significance of correlation coefficients were used.

Results indicated a statistically significant difference between prediction formulas at the .01 level, indicating that the four formulas

did not agree in their prediction of the difficulty of the letters used. When comparing the predictor formulas against mean word error scores, the Hillestad (23) formula had the only significant correlation with a score of .6201, which is significant at the .01 level. When compared with the mean word error score of all letters, the Hillestad, Mellinger (43), and Uthe (72) prediction formulas were all significant at the .01 level for reliability, although the correlation for the Uthe formula was negative. None of the other formulas obtained a significant correlation when compared to syllabic intensity indicating that there was little agreement in order of difficulty as computed by syllabic intensity and the other three predictive formulas. Henrie (20) concluded that of the four, the Hillestad shorthand difficulty prediction formula was the most valid and reliable. He further claimed that although the syllabic intensity and Mellinger prediction formulas were easier to compute, the extra time and effort required by the Hillestad formula would seem to insure greater accuracy.

Henrie also noted that the two most valid prediction formulas contained more than one variable, and both included the number of words beyond the first 1,500 most frequently used words on the Silverthorn (63) list indicating that this item must be a very significant factor in shorthand difficulty prediction. He also recommended that because of the great differences in the components of the formulas he used, further research in the area of high frequency levels be done and that more study was needed to compare the high frequency words, including brief forms, and words beyond the first 1,500 most frequently used words as well as the word levels in between the first 100 and the first 1,500 in frequency.

Wedell, 1972

A year later, recognizing that there was little, if any, agreement on the factors that contribute to the difficulty of shorthand dictation materials and finding that recent research indicated that word frequency appeared to be a more accurate measure of difficulty than other factors, Wedell (73) conducted a study to determine the relationship of syllabic intensity, word frequency, and shorthand stroke intensity to the difficulty of shorthand dictation materials.

To test his hypotheses, he composed two similar 240 standard word letters which had a constant syllabic intensity between 1.5 and 1.55 and the same percentages of words for groupings found in the Perry (50) study. These controls were maintained for each minute as well as for the three minutes of both takes. The average stroke intensity was allowed to vary, however, and was 2.35 for letter A and 2.75 for letter B. Both letters were dictated at 80 words per minute by the individual classroom teachers to 206 students in 15 post-secondary schools in North Dakota, South Dakota, and Minnesota who were certified by the teachers to be writing at least 80 words per minute but not 100 words per minute. A mean error per word score was developed from analysis of 17 variables including five related to the shorthand stroke variable.

An analysis of mean errors per word of words containing one through five syllables showed a trend for the number of errors to increase as the number of syllables increased on letters A and B, both when considered individually and when considered together. This trend was also evident when word frequency was considered with the error per word increasing as the usage decreased. The mean error per word also tended to increase as

the number of shorthand strokes increased. Positive correlations ranging from .44 to .68 were found for all three factors and a significant difference in difficulty was found in the two letters as measured by stroke intensity of 2.35 and 2.75.

To determine the most significant of the three main variables, a Setwise Multiple Regression procedure was used. The syllabic intensity factors were the first set to drop and the word frequency factors dropped second. Stroke intensity remained as the most significant set with a correlation coefficient of .43.

Conclusions reached indicated that while shorthand stroke intensity is definitely a factor in the difficulty of dictation material, neither it nor syllabic intensity nor word frequency should be used as a single measure in determining difficulty. Wedell (73, p. 58) concluded by indicating that "it is conceivable that factors other than syllabic intensity, word frequency, and shorthand stroke intensity contribute to the difficulty of dictation materials."

#### Pullis, 1974-1975

Pullis (54, 53) conducted two independent studies dealing with copy difficulty a year apart. Using a similar design and procedure for both studies, he attempted to test first whether the triple control of percent high frequency words, average word length, and overall syllabic intensity used in measuring the difficulty of typewriting materials would also indicate the difficulty of shorthand dictation materials; and second, whether published dictation material with the same syllabic intensity and high frequency words were of comparable difficulty.

Using shorthand transcription errors as the criterion, he selected

three 5-minute takes of exactly 700 syllables for each study. The first study used tests that were different from each other on the three critical variables. Average word length was measured by typewriting strokes and percent of high frequency words was determined by the number of words within the most frequently used 1,500. The syllabic intensity of the three letters was 1.31, 1.43, and 1.54; average word length was 5.01, 5.37, and 5.80; and the high frequency word index was 91.47, 86.65, and 83.51 percent. In the second study, the takes were chosen on the basis of their similarity having syllabic intensities of 1.5, 1.51, and 1.52 and high frequency word indexes of 85, 80, and 82 percent respectively.

Thirty-two students were involved in the first study and 29 students were tested in the second. All were college level. For both studies, students were randomly assigned to one of three classrooms and given the three tests in random order on three consecutive days. Paired t-tests were used to test for differences between each combination of the three sets of takes in both studies.

Expecting the three tests in the first study to be easy, average, and hard in that order, Pullis (54) found test two to be significantly easier than test one at the .01 level. Using error scores as the criterion, the tests in the first study ranked in difficult, easy, average order. Pullis concluded that the triple control used to measure difficulty of typewriting material did not accurately measure the difficulty of shorthand dictation material. He also found significant differences in difficulty in published takes marked at the same rate and reported that knowing the value of one of the variables did not aid in predicting the value of either of the other two.

Because the three takes in the second study had similar syllabic



intensities and percentages of the first 1,500 high frequency words, the assumption was that student performance would be comparable. However, the results showed respective mean shorthand transcription errors of 15, 32, and 41 for the three tests, with the difference between Test 1 and 2 and between Test 1 and 3 being significant at the .01 level.

Again Pullis (53) concluded that there were significant differences in difficulty of published shorthand takes marked at the same rate, stating that takes with similar syllabic intensities and percent of high frequency words were not necessarily of comparable difficulty.

In both studies, Pullis (53, 54) cautioned that changes in student performance may be reflecting differences in difficulty of the shorthand dictation material rather than changes in the students' shorthand skill.

#### Nickerson, 1977

In another attempt to measure the difficulty of published dictation tests, Nickerson (47) developed a business vocabulary index based on Perry's (50) list of the 5,000 most frequently used words. She also sought to determine the extent to which contemporary business vocabulary was used in five-minute Gregg dictation materials and whether a significant difference in difficulty would be evident in those takes when measured by number of transcription errors.

Using the premise that the most frequently used words in business correspondence would be the easiest to transcribe because they would have been automatized through frequent use, she weighted each word in the Perry list by its corresponding frequency and arrived at an index of 183.37 for the list. Similar vocabulary indexes were then determined for the 60 five-minute published dictation takes at 100 words per minute

included in Previewed Dictation, Progressive Dictation, and Speed Dictation. The takes were then arranged in rank order by vocabulary index; and the six with the highest index were classified "easy," the six middle ones were classified "average," and the six with the lowest vocabulary indexes were classified "hard."

The 18 takes were given by tape at the rate of two per week during the last nine weeks of the quarter to 41 students in fifth quarter shorthand at the college level in an alternating order of hard, average, easy.

Results of a three-factor analysis of variance and Duncan's multiple range test revealed that although transcription error scores on the hard level test differed significantly from those on the easy and average level tests, the vocabulary index failed to identify three distinctly different levels of difficulty in the 18 tests used. Further findings showed that even though syllabic intensities were similar, when measured by mean transcription error scores, the published dictation materials were not of comparable difficulty. Commenting on this result, Nickerson (47) stated:

On the basis of this finding, it is reasonable to conclude that some device or methodology should be developed which will provide a more reliable basis for assessing relative difficulty than the syllabic intensity measure. It is apparent that students' performance on a given take is to a large extent a measure of the degree of difficulty of the particular take (P. 93).

Other findings suggested that the absolute number of transcription errors made on a test could not be used as an indication of the difficulty of a test since those errors were a function of the interaction between a given student's level of shorthand skill and the difficulty of the test. She also found that although most of the takes had vocabulary distributions similar to those identified by Perry (50), 35 percent of

the takes analyzed varied significantly from the vocabulary levels in Perry's (50) study.

Nickerson (47) recommended further study to find a way to determine copy difficulty for shorthand dictation and testing materials. Suggesting additional work with published materials and controlling critical variables, such as syllabic intensity and the percentages of words within word frequency categories, she cautioned teachers that at present "the dictation achievement attained by their students is to a large extent a selective determination of the copy difficulty of the published dictation materials" (p. 94).

#### Studies Dealing With Related Factors

In addition to studies dealing directly with copy difficulty, the results of a number of other investigations have revealed some conclusions that support the contention that there is a relationship between word frequency and difficulty of copy, both with regard to writing and to transcribing of shorthand outlines.

The assumption is made that frequent practice on a word because of its repeated occurrence in the vocabulary will help students develop the ability to write quickly and correctly. Therefore, studies by Rowe (61), Goetz (18), Klein (34), Young (77), Karaim (32), and Palmer (48) showing that speed of writing is affected more by pauses between words as the student attempts to remember the outline than by the time required to execute the outline itself indicate that the quicker students can recall the proper outline, the faster will be the recording speed.

In addition, the justification for drill on brief forms is primarily based on the philosophy that automatization of the outlines for the words

they represent will increase speed and improve performance. Studies by Kahlstrom (31), Jones (29), Lusk (40), Conerly (5), Minnick (46), and Patrick (49) have shown this to be generally the case.

The "practice makes perfect" rationale also applies to transcription where several studies (Gallenberg (16), Pullis (57), Dortch (10), Klaseus (33), Haggblade (19), Ellingson (12), Jester (28), Iannizzi (27), Tronsue (69), and Hooven (24)) show positive correlations between accuracy of shorthand outlines and accuracy of the finished transcript. Error analyses show that high frequency words have lower percentages of error. Although many incorrect outlines can be transcribed accurately, most researchers conclude that the more correctly the outline is written, the more likely the transcript will be accurate. If repeated exposure to and practice on an outline increases the likelihood that it will be written accurately and thus transcribed correctly, then word frequency has an impact on transcription ability.

Since recording speed and transcription ability are the two components measured in the testing process, it would seem that the more words students could write quickly and accurately, and thus transcribe correctly, the greater would be their achievement on tests.

### Summary

A historical review of the literature related to the difficulty of shorthand dictation materials revealed that over the years many variables have been tested in a variety of attempts to determine a valid and reliable method for measuring copy difficulty.

Some studies have dealt with only one factor, some with several. Characteristics of both the dictation materials and the shorthand system

have been examined for possible clues; and analyses have been made of the shorthand notes, the transcript, and combinations of both. Most studies have dealt with students writing either Gregg Simplified or Gregg Diamond Jubilee, but some investigation and research has been done with other systems, notably Pitman. Speeds used in various experiments have ranged from 50 to 120 words per minute, and both original and published materials have been tested. Two major multiple regression formulas have been developed, but several efforts to validate and/or correlate them have been unsuccessful.

Syllabic intensity of 1.4 which has been used for many years to measure copy difficulty has been repeatedly shown to be not only too low for contemporary correspondence but an inadequate index as well. While studies dealing with shorthand stroke intensity are mixed, studies testing the effect of sentence length, shorthand character count, shorthand word count, actual word count, average word length, the Cloze formula, four readability formulas, punctuation marks, typing stroke intensity, and various components of the shorthand system have been unsuccessful.

The notion that vocabulary level as measured by word frequency has a bearing on copy difficulty has been repeatedly mentioned, and a number of studies have dealt with various approaches to the problem. Using word frequency as identified by several major word list compilations as a factor, significant and/or encouraging results have been obtained in many cases. Some, however, conclude that the high frequency words are most predictive while others claim that low frequency, or words beyond 200, 1,500, or 5,000, are better indicators of copy difficulty. The weight of evidence at present tends to support the contention that low frequency words have the greater impact.

The word frequency/copy difficulty link is further reinforced by the findings of studies done with other aspects of shorthand achievement and transcription indicating that speed of recall and accuracy of outline seem to be greater for high frequency words.

At present, the problem of finding a method for consistently and accurately determining the difficulty of copy used for shorthand dictation purposes is still unresolved. The matter is particularly acute with regard to testing materials, since there is not yet any known method for reliably determining the difficulty of copy used to measure student achievement.

## CHAPTER III

### PROCEDURES

#### Introduction

The procedures used for collecting the data to test the hypotheses in this study are outlined in the following sections:

1. Development of the Test Instrument
2. Selection of Participants
3. Administration of the Tests
4. Evaluation of the Results
5. Analysis of the Data

#### Development of the Test Instrument

Because the major thrust of this study dealt with the concept that test difficulty could be determined and/or changed by varying the percentages of words in individual groups in the Perry (50) word list, a two-fold approach was taken: first, that difficulty levels could be significantly changed by increasing or decreasing the percentages of words in the various categories of the word list, and second that tests could be written that were comparable at a given difficulty level by using similar percentages of words in the various categories.

To test the hypotheses associated with the first objective, the decision was made to use three difficulty levels--easy, average, and

hard. It was arbitrarily determined that the percentages of words in the average level would be those indicated by the Perry (50) study and that the percentages for the easy and hard levels would vary 15 percent each way.

The choice of the 15 percent figure was arrived at partly because that amount is a logical breaking point for the easy level, allowing all words in the easy tests to be within the first 1,500 words and partly because it was intuitively felt that that amount would provide a clear separation of levels without making their differences so great as to be obvious.

To test the hypotheses associated with the second objective, the decision was made to develop two statistically parallel tests at each of the three levels. As a result, six tests were required, two with identical controls at the easy level, two with identical controls at the average level, and two with identical controls at the hard level. It was also determined that all tests would be written especially for this study and that they would deal with business issues to maintain content validity.

The words in the Perry list were then separated into seven major categories with words from 1-100 in category 1, words from 101-500 in category 2, words from 501-1,000 in category 3, words from 1,001-1,500 in category 4, words from 1,501-3,500 in category 5, words from 3,501-5,000 in category 6, and all words over 5,000 in category 7.

The corresponding percentages for frequency of occurrence from the Perry study were then listed beside each of those categories, and tests of average difficulty were identified as those containing the same percentages of words for each category as those from the original study.



These percentages were then increased approximately 15 percent to obtain the easy level figures and decreased approximately 15 percent to obtain the hard level figures.

To insure that the change in percentage to easy and hard from the average figures was proportionately spread through all word categories and would not be lumped at either end, all computations were based on the percent of increase from one category to another. To accomplish this, 15 percent of the percent of increase amount shown for each category in the average level was computed. That amount was then either added to or subtracted from the average level percent of increase to arrive at the percentage of increase for the easy and the percentage of decrease for the hard levels. When this had been done, a logical breaking point for the easy level appeared at the end of the fourth category since 98.2 percent of the words fell within the 1-1,500 range. As a result, a slight adjustment was made in the percentages in the easy level to keep all words within the first four categories.

Decreasing the first four categories by 15 percent resulted in these categories containing 72.66 percent in the hard level tests. The remaining 27.34 percent was distributed over the last three categories in the same proportions as they appeared in the average tests. As a result, a little less than 15 percent was actually added to the last three categories on the hard level tests.

The computations made to determine the percentages of words used in the various categories for the three levels of difficulty are summarized in Table I.

TABLE I  
 PERCENT OF INCREASE FOR WORDS IN EACH  
 OF THREE DIFFICULTY LEVELS

Cate- gory	Word Group	Easy		Average		Hard	
		Cumul. Percent	% of Increase	Cumul. Percent	% of Increase	Cumul. Percent	% of Increase
1	1-100	62.24	62.24	53.4	53.4	45.39	45.39
2	101-500	83.935	21.695	71.9	18.5	61.11	15.72
3	501-1000	94.36	10.425	80.6	8.7	68.50	7.39
4	1001-1500	100.00	$\frac{5.64}{100.00\%}$	85.5	4.9	72.66	4.16
5	1501-3500			93.8	8.3	88.11	15.45
6	3501-5000			96.3	2.5	93.04	4.925
7	Over 5000			100.0	$\frac{3.7}{100.0\%}$	100.00	$\frac{6.965}{100.00\%}$

In order to isolate word frequency as the independent experimental variable to the extent possible, a further determination was made to hold several other factors that might affect copy difficulty constant. These were identified as brief forms, brief form derivatives, syllabic intensity, number of actual words per take, and number of different words per take.

Because of research indicating that takes with high concentrations of brief forms are easier than average (Hillestad (23), Pullis (53), Uthe (72), Patrick (49)) brief forms and their derivatives were held constant for all letters. Based on the findings of Reese and Smith (60) which indicated that these factors comprise 44.73 percent and 3.16

percent respectively of average business correspondence, the percentages of those factors allowed within each one-minute segment and each three-minute test was then computed. These figures were held constant for all six tests and did not vary with the three difficulty levels. In addition, since the brief forms are concentrated in the first 500 words and do not fall equally among all words used, no attempt was made to fit their numbers to the percentages of increase for each category. Table II shows the results of these computations based on 80 and 240 words available for one and three minute segments of the tests.

TABLE II  
NUMBER OF WORDS AND RANGE ALLOWED FOR BRIEF FORMS AND  
BRIEF FORM DERIVATIVES IN ONE- AND THREE-MINUTE  
SEGMENTS OF THE SIX STUDY TESTS

	<u>Brief Forms</u>		<u>Brief Form Derivatives</u>	
	Number of Words	Range	Number of Words	Range
80 words	35.78	34-38	2.53	2-3
240 words	107.35	102-113	7.58	7-8

Even though a number of studies have found the relationship of syllabic intensity to copy difficulty to be weak, the fact that it is still listed on most dictation materials and mentioned frequently in

the literature merits its inclusion. Some controversy exists over what the syllabic intensity of average correspondence should be, and figures cited range from the 1.4 that has been traditionally used to 1.65 as recommended by Perry (50). However, it is generally agreed that 1.4 is too low; and many teachers feel that copy with a syllabic intensity of 1.65 is too difficult. It was also felt by the researcher that a jump from the 1.4 that students are familiar with to one that is 1.6 or more would be confusing to them and affect their overall performance. As a result, a compromise of 1.5 (which was recommended by Wedell (73), West (76), and Mellinger (42)) was chosen for the tests developed for this study.

Syllabic intensity was controlled for each minute of each take as well as for the take as a whole. The summary shown in Table III reveals a range of 1.49 to 1.54 with an overall (rounded) average of 1.5.

TABLE III  
SYLLABIC INTENSITY FOR ONE- AND THREE-MINUTE  
PORTIONS OF SIX STUDY TESTS

	First Minute	Second Minute	Third Minute	Average Total
Easy 1 ( $E_1$ )	1.538	1.525	1.488	1.516
Easy 2 ( $E_2$ )	1.500	1.513	1.513	1.508
Average 1 ( $A_1$ )	1.513	1.488	1.500	1.500
Average 2 ( $A_2$ )	1.538	1.513	1.525	1.525
Hard 1 ( $H_1$ )	1.488	1.500	1.525	1.504
Hard 2 ( $H_2$ )	1.525	1.525	1.512	1.520

The fact that the number of actual words and/or the number of different words may have an impact on copy difficulty has been alluded to in the literature because of the effect they may have on the number of errors possible in each take. Letters with many different words or a greater number of total words can generate more errors than their opposite, even though other factors such as syllabic intensity are held constant. To date, little is known about these variables; but because of the possibility referred to and the fact that error scores were used as the criterion, the number of actual words and the number of different words were held constant to avoid any possible contamination. Therefore, the number of actual words was held constant at 80 for each minute and 240 for all three minutes in each of the six study tests.

To determine the number of different words per test, all letters were keypunched on IBM computer cards with one word per card. Each test was then alphabetically sorted by word on the sorter and a printout was obtained. Following this, a manual count was made of all different words for each test.

Because of limitations imposed by other controls, it became extremely difficult to get the number of different words exactly the same for all tests without distorting the content or destroying the sense of some of the letters. As a result, a 5 percent range was accepted, allowing a variance of 12 words among the tests. As can be seen from Table IV, the actual variance in total number of different words for all six tests was nine with no variance greater than three within difficulty levels.

TABLE IV  
NUMBER OF DIFFERENT WORDS FOR EACH OF SIX STUDY TESTS

Difficulty Level	Number of Different Words
Easy 1	151
Easy 2	148
Average 1	154
Average 2	153
Hard 1	157
Hard 2	155

Internal consistency was also maintained as far as possible by having each minute of each three-minute test comparable to both of the other two minutes in terms of percentages of words for vocabulary levels and factors held constant. As a result, all controls for word frequency percentages, brief forms, brief form derivatives, and number of actual words were applied to each one-minute segment within each test as well as to the overall three-minute span.

In keeping with the Perry (50) study, no digital numbers and no proper names were used. Guidelines for counting hyphenated and compound words were also followed. However, because of the variety of controls involved, it was not possible to exclude the salutation and complimentary close from the counts as was done in the Perry study.

A form (See Appendix A) was developed to record and tabulate the number of words needed and/or used in each frequency category, the

number of brief forms and their derivatives, and the number of actual words used as the letters were being written.

As the writing progressed, however, it became evident that some flexibility in the number of words used for the various frequency categories was necessary. As a result, a 5 percent plus or minus variance was allowed based on the total number of words for a given category. To determine the number of words within the 5 percent over or under range, 5 percent of each of the amounts apportioned from the 80 and 240 words was computed and both added to and subtracted from the number of words allowed in each frequency category to give the upper and lower limits. Those figures are shown in Table V.

As each minute of dictation was completed, a color coding system was used to identify the frequency category of each word so that re-checks and final counts could be made. When all three minutes of a test were finished, an over-all tally was made and syllabic intensity calculated. A complete review of all tests, counts, and computations was then made by an assistant to verify all figures and the accuracy of the coding.

All necessary adjustments were then made to insure that all factors were within the specified limits and ranges for each minute and each test at every difficulty level. (See Appendix B for completed tests.)

To allow for simultaneous administration of the tests and to control for any variances that might occur in dictators and/or dictating styles, all six tests were marked for dictation at 80 words per minute and recorded on individual master cassette tapes by the researcher with the assistance of professional personnel and equipment in the sound studio at the Oklahoma State University Audio-Visual Center.

TABLE V  
 NUMBER OF WORDS ALLOWED IN EACH WORD GROUP  
 FOR THREE DIFFICULTY LEVELS

Group	Easy				Average				Hard			
	80 Words	Range	240 Words	Range	80 Words	Range	240 Words	Range	80 Words	Range	240 Words	Range
1-100	49.8	47-52	149.4	142-157	42.7	41-45	128.2	122-135	36.3	34-38	108.9	103-114
101-500	17.4	16-18	52.1	49-55	14.8	14-16	44.4	42-47	12.6	12-13	37.7	36-40
501-1000	8.3	8-9	25.0	24-26	7.0	7	20.9	20-22	5.9	6	17.7	17-19
1001-1500	4.5	4-5	13.5	13-14	4.0	4	11.8	11-12	3.3	3	10.0	9-10
1501-3500					6.7	6-7	20.0	19-21	12.4	12-13	37.1	35-39
3501-5000					2.0	2	6.0	6	3.9	4	11.8	11-12
Over 5000					3.0	3	8.9	8-9	5.6	5-6	16.7	16-18

Note: Some totals may vary slightly from 80 and 240 because of rounding.



The format for each test tape included information for the students, a one-minute warmup at 100 words per minute, and one three-minute test letter at 80 words per minute. (See Appendix C for complete transcript.)

### Selection of Participants

Because of the gradual migration of advanced shorthand classes to the post-secondary level and because of the availability of classes with larger enrollments at that level, participants chosen for this study consisted of students enrolled in post-secondary institutions in the State of Oklahoma during Fall Semester, 1978. Only students studying Gregg shorthand Diamond Jubilee Series and in teacher-directed classes were included. All classes met for 50 minutes per day and on a four- or five-day per week schedule.

A further limitation was made to include only those students writing at least 80 but less than 100 words per minute on a three-minute test dictation. This decision was made to avoid confounding the study with possible differences that might occur at varying skill levels and to equate students as nearly as possible on the basis of shorthand ability.

Telephone contacts were made with all post-secondary schools in the State of Oklahoma to ascertain whether classes working at the 80 to 100 words per minute range were being offered and the procedure used for conducting the class. Based on this initial survey, several schools were eliminated because they did not offer a class for that skill level during Fall Semester. Two others were eliminated because their teaching procedure did not match the criteria established, and one school was eliminated because of lack of support from the administration for

participation in research studies of this nature. Three schools had extremely low enrollments with only three or four students studying in individualized situations at that range. Those schools were also eliminated.

Teachers working with the remaining eight classes at six schools were asked to participate, and all agreed.

Selection of individual students within the classes and verification of their writing speeds was made by means of a two-day preliminary test administered the week before the tests for the study were given. The first preliminary test was given at 80 words per minute and the second was given at 100 words per minute. Since no previously validated tests were available, two tests at each speed were taken from a current tests and awards booklet published by Gregg McGraw-Hill. Because the syllabic intensity of the test material for this study was placed at 1.5, however, the items used for the preliminary tests were recounted and the syllabic intensity was changed from 1.4 to 1.5 to make the preliminary tests equal to the tests in this study in that regard. In doing so, it was necessary to add a few extra words to the published material in order to maintain the three-minute length.

The preliminary tests and instructions were then recorded on master cassette tapes by the researcher at the Oklahoma State University Audio-Visual Center, and duplicate copies were made for the teachers participating in the study.

The taped tests were then given by the individual teachers on two consecutive class periods the week prior to the study. All transcripts were hand scored by the researcher and an assistant using the same criteria outlined for the tests in this study. A listing of all

students who took the preliminary tests together with their error scores was then made. Students were required to have 12 or fewer errors on the 80 words per minute test and 15 or fewer errors on the 100 words per minute tests (95 percent accuracy) to receive a passing score. Only those students who took both preliminary tests and who passed at least one of the 80 takes but neither of the two 100 takes were included in the study. A further restriction on participants required that all students who qualified for inclusion based on the preliminary tests take all six of the study tests. Make-up tests were allowed provided they were given within a week of being missed. Students for whom complete sets of data were not available were not included in the study analysis.

A final count indicated that 193 students took the preliminary tests, 119 students qualified, and 106 students were included in the study. Table VI indicates the number of students participating from each of the eight classes.

TABLE VI  
LATIN SQUARE, GROUP, AND NUMBER OF STUDENTS  
PARTICIPATING FROM EACH CLASS

Class	Number Participating	Group	Latin Square
1	13	1	A & B
2	18	2	A & B
3	21	3	A & B
4	9	4	C & D
5	13	5	C & D
6	18	6	C & D
7	8	6	C & D
8	6	6	C & D

### Administration of the Tests

To acquaint teachers with the materials, answer questions, and insure uniformity of procedure, a personal visit was made by the researcher during the last week of October to each school involved in the study. Each teacher was given a set of written guidelines (See Appendix D), several large mailing envelopes, and address labels. Each teacher was also given the two tapes containing the preliminary tests for administration the following week.

To help students and teachers become acquainted with the testing situation, the researcher's voice, and the procedure to be followed in the study, the format of the preliminary tests followed closely that of the tests used for the study. In the introduction given on the tape (See Appendix E for a complete transcript), students were encouraged to write and transcribe every word possible, whether they felt their paper would receive a passing score or not. They were given instructions regarding the format to be followed for the transcript and told to put their name, the name of their school, the current date, and a code number that identified the take in the upper right hand corner of their transcript. They were not told that these were qualifying tests or that they were participating in an experiment. To insure maximum effort, teachers had previously told students that all tests passed would be recorded and count toward their grades.

On the first preliminary tape, the introductory instructions were followed by a one-minute warmup at 100 words per minute and then two, three-minute tests at 80 words per minute. Students were told to transcribe either or both of the takes depending on the quality of

their notes and were given the remainder of the class period, amounting to 35 or 40 minutes, to do so. The transcription period was not timed more closely than indicated in order to allow students maximum time to transcribe all they could. At the end of the class period, teachers collected both the transcript and the shorthand notes and sent them to the researcher for scoring.

The second day of the preliminary tests followed the same procedure as the first, except that the warmup was at 120 words per minute and the two three-minute tests were dictated at 100 words per minute.

When the number of students available for the study had been determined, they were separated by class into six groups. To accomplish this, the first five classes were assigned to groups 1 through 5, and the remaining three classes were lumped together to form group 6 (See Table VI). This grouping facilitated the use of the Latin Square design which provided a systematic way to account for three sources of variability in the data which were identified as (1) differences in ability levels among classes, (2) learning that may have occurred during the testing sequence, and (3) the difficulty level of the test.

The four random 3 x 3 Latin Squares which were used to account for the three variables described above are shown in Figure 1.

As can be seen, the first version of the easy, average, and hard tests were randomly assigned to groups 1, 2, and 3 the first week (Square A), and the second version of each of the three levels was randomly assigned to the same three groups the second week (Square B).

Groups 4, 5, and 6 were randomly assigned the second version of the easy, average, and hard test the first week (Square C) and the first version of the three tests the second week (Square D). The use of the

Latin Square insured that each difficulty level occurred first, second, and third in each grouping and that each version of the test occurred both first and second. This arrangement also assured that all students took all tests at all difficulty levels.

FIGURE 1

SQUARE A			SQUARE C		
Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
H <sub>1</sub>	E <sub>1</sub>	A <sub>1</sub>	H <sub>2</sub>	E <sub>2</sub>	A <sub>2</sub>
A <sub>1</sub>	H <sub>1</sub>	E <sub>1</sub>	E <sub>2</sub>	A <sub>2</sub>	H <sub>2</sub>
E <sub>1</sub>	A <sub>1</sub>	H <sub>1</sub>	A <sub>2</sub>	H <sub>2</sub>	E <sub>2</sub>
SQUARE B			SQUARE D		
E <sub>2</sub>	A <sub>2</sub>	H <sub>2</sub>	A <sub>1</sub>	H <sub>1</sub>	E <sub>1</sub>
A <sub>2</sub>	H <sub>2</sub>	E <sub>2</sub>	H <sub>1</sub>	E <sub>1</sub>	A <sub>1</sub>
H <sub>2</sub>	E <sub>2</sub>	A <sub>2</sub>	E <sub>1</sub>	A <sub>1</sub>	H <sub>1</sub>

Figure 1. Layout of Experimental Design

Teachers were then provided with a set of six individual cassette tapes (duplicated from the master test tapes) which contained the controlled tests for this study. Each set of these tapes was coded and labeled in the order in which tests were to be given in each particular group. As a result, some students received the tests ranging from easy to hard, others received them from hard to easy, and still others received average first followed by either easy or hard. In addition, some students received the second version of the tests first and the first version second, and others received the first version first and the second version second.

All tests were given during the second and third weeks of November, 1978. Students received one test per day on three consecutive days of each week, avoiding Monday and Friday wherever possible. Contact with the teachers was maintained throughout the testing period by telephone, mail, and personal visit. Teachers were instructed to say as little as possible about the tests or their purpose and to answer any questions in a very general way. Normal classroom routine was maintained as closely as possible, and no previews or advance discussion of the tests were permitted. Teachers were instructed to play the tape as soon as possible after the beginning of the class, monitor the transcription, and collect the transcripts and shorthand notes at the end of the class period. Transcripts were then given or mailed to the researcher for hand scoring. Teachers did not know the sequence or difficulty level of the tests being administered to their students in advance. When all tests had been given, the teachers also returned the cassette tapes to the researcher.

#### Evaluation of the Data

For consistency in procedure, the typewritten transcript of each test was hand scored by the researcher and one assistant. For purposes of this study, only shorthand errors on the transcript were circled and counted and only one error per word was counted. No spelling, grammar, punctuation, capitalization, paragraphing, or typographical errors were considered, nor were errors in the shorthand notes counted. Individual error scores for each transcript were based on total number of words omitted, added, mistranscribed, or that otherwise varied from the dictation. In the few cases where students did not finish transcribing, an

error was counted for each missing word. The total number of circled words was then tallied and shown at the top of each transcript as the raw transcription error score for that paper.

Score sheets listing students in each group (See Appendix F) were prepared, and a record was kept of each student's raw transcription error score for each test as the transcript was received and scored. A total of 636 transcripts were evaluated representing six transcripts from each of the 106 students participating.

### Analysis of the Data

When raw transcription error scores had been determined for all students on all tests (See Appendix G), the raw mean transcription error score for each group on each of the six tests was computed (See Appendix H). To determine if the two similar tests at each of the three levels of difficulty were statistically equivalent, a two-way analysis of variance procedure was used, and a separate analysis was run for each of the eight classes.

A weighted least squares analysis was used instead of an ordinary analysis of variance to determine if the difficulty of the tests at each of the three levels was statistically different because the variability between test scores depended on the level of difficulty and because there were unequal numbers of students in each group.

Using means from the cells of the four random 3 x 3 Latin Squares, adjusted mean transcription error scores and estimates of the standard errors of the difference between means were computed by a consulting statistician using a computer program for the weighted least squares procedure. The adjusted mean transcription error scores were then used to



determine differences in mean error scores between the average and easy, the hard and average, and the hard and easy difficulty levels. Z scores were used to test for significance of the differences obtained.

### Summary

In summary, the procedures used to test the hypotheses of this study were concerned with the development and testing of six three-minute dictation tests. Two statistically comparable tests were written at each of three difficulty levels based on percentages of words in seven frequency categories from the Perry (50) list. These tests were then recorded on cassette tapes and given in random order to a group of 106 students writing at least 80 but less than 100 words per minute in several post-secondary schools in Oklahoma during November of 1978.

Transcripts were hand scored by the researcher, and a total shorthand error score was recorded for each student on each test. A two-way analysis of variance was used to determine if the two tests at each of the three levels of difficulty were statistically equivalent. The weighted least squares analysis, a comparison of adjusted mean transcription error scores, and the computation of Z scores were used to determine if the tests at each of the three difficulty levels were statistically different from each other.

## CHAPTER IV

### FINDINGS

#### Introduction

The findings of the study to determine the effect of using both actual and varied percentages of words according to their frequency in normal business correspondence on copy difficulty for shorthand testing materials are reported in three sections:

1. Determination of Equivalency of Tests at Same Difficulty Level
2. Comparison of Three Difficulty Levels
3. Supporting Evidence

The results are based on the data obtained from the administration of six specially constructed shorthand dictation tests at three levels of difficulty to 106 students writing at least 80 but not 100 words per minute in Oklahoma during Fall Semester, 1978. All computations were based on intra-person variances, or those occurring between and among one student's scores on each of the six tests. Since actual achievement, or number of errors, on the easy, average, and hard level tests is relative depending on the skill level of each student, inter-person variance, or comparisons among students, classes, or groups, were not made. These findings, therefore, reflect average differences in difficulty, and not absolute scores, which could be expected between the two tests at each level and between comparisons made among the three

levels of difficulty by students writing at least 80 but not 100 words per minute.

#### Determination of Equivalency of Tests at Same Difficulty Levels

When all students had completed all six tests, a raw mean transcription error score was computed for each of the eight classes on each test (See Appendix H). The two-way analysis of variance was used to evaluate the evidence against the hypotheses that the two similar tests at each level were equivalent. A separate analysis was run for each of the eight classes at each of the three levels of difficulty using the raw mean transcription error scores of both versions of the tests at each level. Somewhat contrary to the typical case for testing null hypotheses where high levels of significant differences in means are used to indicate the degree to which tests differ from each other, the intent of this analysis was to show the lack of difference in the two tests. Therefore, in this particular case, nonsignificant differences are of more value than significant ones to indicate that relationship.

As shown in Table VII, 23 of the 24 tests run were nonsignificant with only one being significant at the .01 level. One of 24 is well within the range of probability that a Type I error may have resulted, and since the remaining 23 nonsignificant F values present extremely strong evidence that the two tests at each difficulty level are equivalent, the null hypotheses 1, 2, and 3 (page 9) indicating that there would be no significant difference in difficulty between the two easy, the two average, and the two hard tests as measured by mean transcription error scores was not rejected.

TABLE VII  
 TWO-WAY ANALYSIS OF VARIANCE COMPARISONS  
 OF TWO EQUIVALENT TESTS AT  
 THREE DIFFICULTY LEVELS

Class	Degrees of Freedom	Computed F Value			F Value Required at .01 Level
		Easy	Average	Hard	
1	1 12	1.6	<1	<1	9.33
2	1 17	<1	<1	<1	8.40
3	1 20	<1	1.63	8.727*	8.10
4	1 8	3.58	3.448	<1	11.26
5	1 12	2.78	1.608	<1	9.33
6	1 17	<1	<1	2.8568	8.40
7	1 7	<1	<1	<1	12.25
8	1 5	<1	<1	<1	16.26

\*Significant at the .01 level

#### Comparison of Three Difficulty Levels

Four random 3 x 3 Latin Squares had previously been constructed to account for any variability caused by differences in group ability, order of testing, and the difficulty level of the tests. Raw mean transcription error scores were then computed for each test taken by each group in every cell in all four Latin Squares. A weighted least

squares procedure was then used to weight these scores to adjust for any effect attributable to unequal group sizes and variability which was dependent on the difficulty level of the test. These adjusted mean transcription error scores represented estimates of the differences in transcription error scores among the easy, average, and hard difficulty levels for each of the six groups and were used to determine the differences in average difficulty, as measured by transcription error scores, for the three levels of difficulty. These data are summarized in Table VIII.

Comparing the two sets of data shows that there is little difference between the raw mean transcription error scores and the adjusted mean transcription error scores. The small discrepancy in the two scores shows that the model used was, in fact, correct. Even though the difference is not great, the weighting helps to avoid some of the loss of sensitivity, and the quality of the answer is higher and more reliable when scores are adjusted in this manner.

To determine the differences between the overall adjusted mean transcription error scores for the three comparisons indicated by the hypotheses 4, 5, and 6 (page 9), appropriate computations were made, and the results are shown in Table IX on page 76.

As can be seen, groups 1, 2, and 3 made an average of 9.52 and 6.55 fewer errors on the first and second versions of the easy test than on the first and second versions of the average test. Those same three groups made an average of 22.83 and 19.64 fewer errors on the two versions of the average test than on the two versions of the hard test. They also averaged 32.35 and 26.19 fewer errors on the easy versions than on the hard versions of the two tests.

TABLE VIII

SUMMARY OF RAW AND ADJUSTED MEAN TRANSCRIPTION ERROR SCORES  
FOR FOUR LATIN SQUARES

Group	Square	Test Order	Raw Mean Transcription Error Scores			Raw Overall Mean for Group	Adjusted Mean Transcription Error Scores			Adjusted Overall Mean for Group
			Easy	Average	Hard		Easy	Average	Hard	
1	A	HAE1	11.31	25.46	50.85	29.21	13.60	22.64	49.39	28.54
2	A	EHA1	3.83	9.00	25.06	12.63	2.47	8.55	30.90	13.97
3	A	AEH1	<u>10.05</u>	<u>21.10</u>	<u>46.71</u>	25.95	<u>9.79</u>	<u>23.23</u>	<u>42.61</u>	25.21
		$\bar{x}$	8.40	18.52	40.87		8.62	18.14	40.97	
1	B	EAH2	15.38	22.92	46.92	28.41	16.91	21.75	44.28	27.65
2	B	AHE2	4.00	9.61	22.83	12.15	3.44	8.80	26.72	12.99
3	B	HEA2	<u>9.91</u>	<u>17.48</u>	<u>39.05</u>	22.15	<u>9.45</u>	<u>18.90</u>	<u>37.35</u>	21.90
		$\bar{x}$	9.76	16.67	36.27		9.93	16.48	36.12	
4	C	HEA2	4.00	18.89	33.44	18.78	7.00	9.31	45.36	20.56
5	C	EAH2	13.77	16.38	35.85	22.00	9.35	19.89	43.16	24.13
6	C	AHE2	<u>11.13</u>	<u>25.91</u>	<u>65.00</u>	34.01	<u>12.08</u>	<u>27.18</u>	<u>58.68</u>	32.65
		$\bar{x}$	9.63	20.39	44.76		9.48	18.79	49.07	
4	D	AHE1	7.11	14.78	37.00	19.63	8.21	10.13	44.03	20.79
5	D	HEA1	8.85	18.69	40.00	22.51	6.71	20.76	42.68	23.38
6	D	EAH1	<u>11.41</u>	<u>25.41</u>	<u>63.00</u>	33.27	<u>11.97</u>	<u>25.87</u>	<u>59.93</u>	32.59
		$\bar{x}$	9.12	19.63	46.67		8.96	18.92	48.88	

$\bar{x}$  = Overall adjusted mean transcription error score for each difficulty level

TABLE IX  
 AVERAGE DIFFERENCES BETWEEN OVERALL ADJUSTED  
 MEAN TRANSCRIPTION ERROR SCORES FOR THREE  
 DIFFICULTY LEVEL COMPARISONS

Difficulty Level	Overall Adjusted Mean Transcription Error Scores			
	Groups 1, 2, and 3		Groups 4, 5, and 6	
	Square A (Version 1)	Square B (Version 2)	Square C (Version 2)	Square D (Version 1)
Average	18.14	16.48	18.79	18.92
Easy	- 8.62	- 9.93	- 9.48	- 8.96
Difference	<u>9.52</u>	<u>6.55</u>	<u>9.31</u>	<u>9.96</u>
Hard	40.97	36.12	49.07	48.88
Average	-18.14	-16.48	-18.79	-18.92
Difference	<u>22.83</u>	<u>19.64</u>	<u>30.28</u>	<u>29.96</u>
Hard	40.97	36.12	49.07	48.88
Easy	- 8.62	- 9.93	- 9.48	- 8.96
Difference	<u>32.35</u>	<u>26.19</u>	<u>39.59</u>	<u>39.92</u>

Groups 4, 5, and 6 averaged 9.31 and 9.96 fewer errors on the two versions of the easy test than on the corresponding versions of the average test. They made 30.28 and 29.96 fewer errors on the two versions of the average test than on the two versions of the hard test. Finally, they made an average of 39.59 and 39.92 fewer errors on the easy versions than on the two hard versions of the tests.

To determine the significance of the differences shown in Table IX, Z scores were computed for each comparison. Using a ratio of the computed average differences in transcription error scores and the standard

error of the difference between means (See Appendix I), the Z scores shown in Table X were obtained. The fact that all scores are considerably higher than +3 standard deviations from the mean indicates their significance.

TABLE X  
Z SCORES FOR THREE DIFFICULTY LEVEL COMPARISONS

Difficulty Level	Square A	Square B	Square C	Square D
Average - Easy	6.46	4.45	5.89	6.29
Hard - Average	11.00	9.46	14.10	13.95
Hard - Easy	16.42	13.29	19.17	19.33

The probability of obtaining the Z scores shown is less than .001 if the null hypotheses being tested were true. As a result, it was determined that there was very strong evidence against the null hypotheses and that there were, in fact, highly significant differences in student's mean transcription error scores on the easy versus the average, the average versus the hard, and the easy versus the hard level tests. Hypotheses 4, 5, and 6 were therefore rejected.



### Supporting Evidence

Additional evidence that the two versions of the tests were of comparable difficulty was obtained by determining the relative achievement of groups within squares. To accomplish this, computations were made to determine mean transcription error differences of each group compared to the other two groups within the same square. It was expected that the average performance of each group above or below the overall adjusted mean would be approximately the same for each of the two weeks.

To estimate the average number of errors each group averaged above or below the overall adjusted mean error score of each square, the total number of mean transcription errors per group was subtracted from the overall square mean. This figure measured the group average performance on the three tests in each square. These results are summarized in Table XI, which shows the relative achievement of each group within the two squares in which their scores fall. These scores range from -8.60 to +7.0 overall; but as can be seen, scores for each group from Week 1 to Week 2 do not vary greatly indicating that the performance of each group was quite consistent on the two sets of tests.

Additional verification of the findings was obtained with a two-way analysis of variance using grand mean transcription error scores for all six tests. The very low F value for the difference between the equivalent tests and the extremely high F value for the differences in difficulty level shown in Table XII tend to confirm the results already presented and verify that there was no significant difference between the equivalent tests but great differences among the difficulty levels.

TABLE XI  
RELATIVE ACHIEVEMENT OF GROUPS WITHIN SQUARES

Group	Square	Average number of errors above or below overall adjusted mean (Week 1)	Square	Average number of errors above or below overall adjusted mean (Week 2)
1	A	5.97	B	6.80
2	A	-8.60	B	-7.86
3	A	2.63	B	1.06
4	C	-5.22	D	-4.80
5	C	-1.65	D	-2.20
6	C	6.87	D	7.00

TABLE XII  
TWO-WAY ANALYSIS OF VARIANCE FOR TWO EQUIVALENT TESTS  
AND THREE DIFFICULTY LEVELS  
USING GRAND MEANS

Variable	Degress of Freedom	Computed F Value	Required at .01 Level
Test	1	0.76	98.49
Difficulty Level	2	491.58	98.49

### Summary

The findings of this study are based on an analysis of the mean transcription error scores obtained from the administration of six three-minute tests designed to determine whether two equivalent tests at each of three levels were equal in difficulty. The data were also analyzed to determine whether there were significant differences in adjusted mean transcription error scores on the easy versus the average, the average versus the hard, and the easy versus the hard difficulty level tests

Based on F values obtained from a two-way analysis of variance of the data, there was no statistically significant difference in the difficulty of the two equivalent tests at each level.

Results obtained by using a weighted least squares analysis on four random 3 x 3 Latin Squares to obtain adjusted mean transcription error scores, the comparison of adjusted mean transcription error scores among difficulty levels, and the computation of Z scores as tests of significance revealed that there were highly significant differences in the adjusted mean transcription error scores for all three difficulty level comparisons made.

## CHAPTER V

### SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

#### Summary

Teachers and researchers have long recognized the need for a valid and reliable method for determining the difficulty of copy used for measuring shorthand achievement. Syllabic intensity, which has been used for this purpose for many years, has been repeatedly discredited by studies showing its low predictive power and correlation as measured by both shorthand and transcription error scores. Of the variety of alternatives discussed as possible determinants of copy difficulty, word frequency has been most often mentioned; and many suggestions have been made that further research be done with this factor to discover what, if any, relationship exists. The intent of this study, therefore, was to attempt to identify criteria based on word frequency that could be used to develop copy of known difficulty for shorthand testing materials.

Based on the assumption that Perry's (50) listing of most frequently used words is representative of average business correspondence, answers were sought to the following questions:

1. Is it possible to write shorthand testing materials that are of comparable difficulty by controlling the percentages of words used in various frequency categories while holding brief forms, brief form derivatives, syllabic intensity,

number of different words, and number of actual words constant?

2. Can the difficulty of those tests be varied by changing the percentages of words in the different frequency categories by approximately 15 percent while holding brief forms, brief form derivatives, syllabic intensity, number of different words, and number of actual words constant?

The six null hypotheses listed below were formulated to test at the .01 level whether materials having similar percentages of words from word groupings identified by Perry (50) would be of comparable difficulty and whether the difficulty level of shorthand testing materials could be significantly changed by varying the percentages of different categories of frequently used words.

1. There is no statistically significant difference in average difficulty between two easy tests as measured by mean transcription error scores.
2. There is no statistically significant difference in average difficulty between two average tests as measured by mean transcription error scores.
3. There is no statistically significant difference in average difficulty between two hard tests as measured by mean transcription error scores.
4. There will be no statistically significant difference in student's mean transcription error scores on the easy difficulty level tests and the average difficulty level tests.
5. There will be no statistically significant difference in student's mean transcription error scores on the

average difficulty level tests and the hard difficulty level tests.

6. There will be no statistically significant difference in student's mean transcription error scores on the easy difficulty level tests and the hard difficulty level tests.

A review of the literature indicated that while many studies have attempted to identify factors that could be used to predict and/or measure copy difficulty, no valid and reliable method has yet been found.

Some support for syllabic intensity, which has been traditionally used, was offered by Leslie (36), Elsen (13), Hillestad (23), and Wedel (73), but the greater majority of studies using this variable have found it inconsistent and/or inadequate. Likewise, shorthand stroke intensity and shorthand stroke count were found to be helpful as predictors by two researchers, but not significantly related to difficulty by others. Sentence length, average word length, and several readability formulas were also found to be low in the ability to determine difficulty. Hillestad (23) and Uthe (72) both developed multi-regression formulas for measuring copy difficulty, but subsequent efforts to validate those studies were unsuccessful.

Word frequency and vocabulary indexes based on word lists developed by Horn (25), Silverthorn (63), Perry (50), and Mellinger (43) and been used as factors in a number of studies; and, in general, the conclusions support the concept that there is a relationship between word frequency and copy difficulty. What the interaction is and how it can be used to measure copy difficulty remains to be discovered, however.

Procedures used for collecting the data to test the hypotheses of

this study involved the development of six three-minute shorthand dictation tests that were given at 80 words per minute to 106 post-secondary students in eight classes in Oklahoma during the second and third weeks of November, 1978. The six tests were composed of two equivalent tests at each of three difficulty levels identified as easy, average, and hard.

Tests classified as average contained the same percentages of words in seven major word groups as those identified by Perry (50). Tests classified as easy contained approximately 15 percent more words in the high frequency groups shifted from the low frequency range, and tests classified as hard had an approximately 15 percent greater concentration of words in the low frequency words groups with a corresponding decrease in the high frequency word groups.

All tests were written especially for this study and dealt with business issues. The percentages of brief forms and brief form derivatives, syllabic intensity, the number of different words, and the number of actual words were held constant for all tests. In addition, internal control was maintained by holding percentages of brief forms and brief form derivatives, syllabic intensity, and number of actual words constant for each minute as well as for all three minutes of each test.

As a result, all tests contained 44.73 percent and 3.16 percent respectively of brief forms and brief form derivatives, an average syllabic intensity of 1.5, and 80 words each minute for a total of 240 words in all three minutes.

The six tests were then recorded on tape at the Oklahoma State University Audio-Visual Center. Each tape contained instructions to

the students, a one-minute warmup at 100 words per minute, and one three-minute test at 80 words per minute.

Participants in the study were selected on the basis of their ability to pass at least one of two preliminary three-minute dictation tests at 80 words per minute but neither of two similar tests at 100 words per minute. These tests were given on two consecutive days the week before the study began and were adapted from those published in a current tests and awards booklet by Gregg McGraw-Hill.

To account for variability caused by test order, differences in group ability, and/or the difficulty level of the test, four 3 x 3 Latin Squares were developed. Each test occurred once in every position of first, second, and third; and every student took every test. The two versions of the test were also arranged in opposing order so that each version was given both first and second. The test tapes were then coded and labeled in the order in which they were to be given to each of the groups according to their position in the Latin Square design.

Students in the eight classes were placed in six groups and each group was assigned to a column on the Latin Squares. Since the six tests had been previously randomized using the four 3 x 3 Latin Squares, this assignment determined the order in which each particular group received the tests.

Tests were administered by tape at the rate of three per week for two weeks by the individual teacher according to instructions from the researcher. Students were given one easy, one average, and one hard test in randomized order each week.

All transcripts were hand scored by the researcher counting only shorthand related errors such as additions, mistranscriptions, and



omissions of words as errors. Raw mean transcription error scores were computed for each of the six groups and eight classes on every test.

Nonsignificant F values obtained by a two-way analysis of variance on raw mean transcription error scores for each of the eight classes indicated that there was no significant difference in the two equivalent tests at each difficulty level, and the hypotheses associated with the question regarding the ability to develop tests of equal difficulty were not rejected.

To determine whether there was a significant difference in difficulty among the three levels of tests, the raw mean transcription error scores for each of the six groups was treated with the weighted least squares analysis procedure which adjusted for any variances caused by unequal group sizes and/or variability due to the difficulty level of the test. Analysis of the adjusted mean transcription error scores revealed average differences between levels for all three comparisons that were highly significant at the .001 level when Z scores were computed. As a result, the three hypotheses indicating no significant difference between the easy and the average, the average and the hard, and the easy and the hard difficulty levels were rejected.

### Conclusions

The following conclusions were drawn from the analysis of the data obtained from this study.

1. It is possible to develop shorthand dictation tests of comparable difficulty by using similar percentages of words from specified frequency categories in the Perry list while holding brief forms, brief form derivatives,

syllabic intensity, number of actual words, and number of different words constant.

2. The difficulty of copy used for shorthand dictation tests can be significantly changed by increasing or decreasing the percentages of words in the various frequency categories by approximately 15 percent while holding brief forms, brief form derivatives, syllabic intensity, number of actual words, and number of different words constant.

### Recommendations

Based on the findings of this study, several recommendations for implementation and further research are suggested below:

1. Absolute error scores on shorthand dictation tests are a function of student skill and test difficulty. Therefore, teachers and researchers should measure gains in achievement primarily by changes in scores from one test to another test of comparable difficulty.
2. Those using the results of this study should be aware that the findings are based on mean differences in number of errors between tests of equal difficulty and between tests at three levels of difficulty and are not student scores per se. They are estimates of the differences in the average number of errors students can be expected to make on the various levels of tests regardless of individual student scores. Student scores on tests of equivalent difficulty will be comparable, and students will likely have about 9 more errors on the average than the easy

level tests, about 25 more errors on the hard than the average level tests, and about 35 more errors on the hard than the easy level tests regardless of what their actual scores turn out to be.

3. To improve the validity and reliability of current testing practices, new testing materials based on the criteria identified in this study should be developed and used for all future classroom and research evaluation of student progress and achievement.
4. The difficulty level of existing testing materials should be determined and published wherever possible so that appropriate selection of such tests can be made by those using them.
5. A study should be undertaken to determine whether a variation of less than 15 percent in the word frequency categories will produce results similar to those of the present study.
6. To determine the effect of other variables on copy difficulty, a study in which word frequency is held constant but factors such as brief forms, brief form derivatives, number of actual words, and number of different words is allowed to vary should be done.
7. A study similar to the present one but using other speed ranges such as 60 to 80 words per minute or 100 to 120 words per minute should be done to determine if the results are consistent for all speeds.

8. A study should be conducted to determine whether similar results are obtained with other shorthand systems.
9. It is possible that other, or fewer, word groupings may produce results similar to those found in this study. Therefore, it is suggested that experimentation be undertaken to determine the effect of using different arrangements or groupings of word categories to identify the different levels of difficulty of shorthand dictation material.
10. A study should be conducted to investigate the impact of controlling for each minute as well as for all three minutes of a shorthand dictation test item.
11. Further research should be conducted to identify additional categories of difficulty such as "very easy" and "very hard" and to define more closely the limits of easy, average, and hard used in this study.

## SELECTED BIBLIOGRAPHY

1. Adams, Richard E. "Professional Shorthand Teacher's Corner." Century 21 Reporter (Fall, 1974), p. 8.
2. Baggett, Harry W., Jr. "The Validity of a Measure of the Difficulty of Gregg Shorthand Dictation Materials." (Unpub. Ph.D dissertation, University of Minnesota, 1964.)
3. Bartz, Albert E. Basic Statistical Concepts in Education and the Behavioral Sciences. Minneapolis: Burgess Publishing Company, 1976.
4. Boggess, Violet F. "The Results of Using Dictation Materials of Varying Difficulty for Speed Practice in Shorthand Classes." (Unpub. Ph.D dissertation, The Ohio State University, 1970.)
5. Conerly, Donna Lynn. "The Effect of the Increased Use of Brief Forms and Their Derivatives on the Performance of Students in Beginning and Intermediate College Shorthand." (Unpub. Ed.D dissertation, University of Southern Mississippi, 1975.)
6. Crandall, Lars G. "Word Frequency Applied to Stenography." The Journal of Business Education, 63 (November, 1960), pp. 67-68.
7. Curtin, Rita C. "The Relationship Between Selected Factors and Difficulty of Dictated Material." (Unpub. MA thesis, University of Minnesota, 1958.)
8. Danielson, Harriet A. "The Relationship Between Competency in Shorthand Vocabulary and Achievement in Shorthand Dictation." (Unpub. Ed.D dissertation, Indiana University, 1959.)
9. Dewey, Godfrey. The Relativ (sic) Frequency of English Speech Sounds. Cambridge: Harvard University Press, 1923.
10. Dortch, Richard Neil. "A Study to Determine the Relationship Between Shorthand Outlines Constructed According to Theory and the Accuracy of the Transcript." (Unpub. Ph.D dissertation, The University of Wisconsin-Madison, 1976.)
11. Eldridge, R. C. Six Thousand Common English Words. Niagara Falls: By the Author, 1911.

12. Ellingson, Jean T. "An Analysis of Errors Made in Shorthand Outlines and the Transcript to Determine Critical Nature of Errors." (Unpub. MS thesis, Mankato State College, 1966.)
13. Elsen, Sister Mary Elfrida. "Factors of Difficulty in Shorthand Dictation Material." (Unpub. MA thesis, DePaul University, 1946.)
14. Farmer, Geraldine. "An Experiment to Test the Validity of a Measure of the Difficulty of Shorthand Dictation Materials." (Unpub. Ph.D dissertation, University of Minnesota, 1961.)
15. Flood, Hazel A. "Some Factors in Learning Shorthand and an Analysis of Two Systems of Shorthand." (Unpub. Ph.D dissertation, University of Minnesota, 1958.)
16. Gallenberg, Barbara. "A Study Showing the Relationship Between the Writing of Theoretically Correct Shorthand Outlines to the Transcription Process." (Unpub. MS thesis, University of Wisconsin-Madison, 1975.)
17. Gallion, Leona M. "The Teaching of Shorthand and Transcription." Contributions of Research to Business Education. Ninth Yearbook of the National Business Education Association. Washington, D. C.: National Business Education Association, 1971. pp. 38-47.
18. Goetz, Leo G. "The Relationship Between Symbol Mastery and Selected Dictation Speeds in Gregg Shorthand." (Unpub. Ed.D dissertation, The University of North Dakota, 1969.)
19. Haggblade, Berle. "Factors Affecting Achievement in Shorthand." (Unpub. Ed.D dissertation, University of California-Los Angeles, 1965.)
20. Henrie, Bill S. "A Comparative Analysis of Difficulty Prediction Formulas for Shorthand Dictation Material." (Unpub. Ed.D dissertation, Utah State University, 1971.)
21. Henshall, Joy L. "An Application of Readability Techniques to Prediction of Difficulty Level of Shorthand Dictation Materials." (Unpub. Ph.D dissertation, North Texas State University, 1971.)
22. Hess, Susie J. and Patrick, Alfred L. "A Design for Second-Year Shorthand Instruction." The Balance Sheet, LVII (March, 1976), p. 244-48, 278.
23. Hillestad, Mildred C. "Factors Which Contribute to the Difficulty of Shorthand Dictation Materials." (Unpub. Ph.D dissertation, University of Minnesota, 1966.)

24. Hooven, Jean A. "The Effects of Intensive Practice of a Target Vocabulary in First-Year Gregg Shorthand." (Unpub. Ed.D dissertation, Temple University, 1978.)
25. Horn, Ernest. A Basic Writing Vocabulary, Ten Thousand Words Most Commonly Used in Writing. Iowa City: University of Iowa Monographs in Education, First Series, No. 4, 1926.
26. Horn, Ernest, and Peterson, Thelma. The Basic Vocabulary of Business Letters. New York City: The Gregg Publishing Company, 1943.
27. Iannizzi, Elizabeth. "Transcription and Shorthand Errors Among Elementary and Advanced High School Writers of Simplified and Diamond Jubilee Gregg Shorthand." (Unpub. Ph.D dissertation, New York University, 1967.)
28. Jester, D. D. "A Time Study of the Shorthand Transcription Process." (Unpub. Ph.D dissertation, Northwestern University, 1959.)
29. Jones, Jo Nell. "Testing a Spaced Pattern of Repetition of Gregg Diamond Jubilee Shorthand Brief Forms and Principles at the Community College Level." (Unpub. Ed.D dissertation, The University of Tennessee, 1975.)
30. Kaedig, F. W. German Frequency Word Book, Based on Kaedig's Haufigkeitswörterbuch der deutschen Sprache. New York City: The Macmillan Company, 1928.
31. Kalstrom, Roger. "An Analysis of Errors in Writing Brief Forms in Gregg Shorthand." (Unpub. MS thesis, Mankato State College, 1957.)
32. Karaim, William J. "A Comparison of the Writing Practices of Two Diverse Achievement Groups of Shorthand Students When Writing From Dictation." (Unpub. Ed.D dissertation, University of North Dakota, 1968.)
33. Klaseus, R. C. "An Analysis of Some of the Factors that Contribute to the Difficulty of Transcription Materials in Gregg Shorthand, Diamond Jubilee Series." (Unpub. MS thesis, Mankato State College, 1964.)
34. Klein, Abraham E. "Variations in the Speed of Writing of Symbol Combinations in Gregg Shorthand." National Business Education Quarterly, 18 (October, 1949), pp. 35-44.
35. Leslie, Louis A. Methods of Teaching Transcription. New York: Gregg Publishing Division, McGraw-Hill Book Company, Inc., 1949.
36. Leslie, Louis A. "Providing Controlled Dictation." American Shorthand Teacher, XII (September, 1931), pp. 21-22.

37. Leslie, Louis A. "The Difficulty of Shorthand Dictation Material." Business Education World, XXVII (September, October, 1947), pp. 13-16, 35-37.
38. Leslie, Louis A. "The Standard Word in Shorthand Dictation." Business Education World, XXIV (January, 1944), pp. 280-83.
39. Leslie, Louis A. The Story of Gregg Shorthand. New York City: McGraw-Hill Book Company, 1964.
40. Lusk, Norman M. "A Study of the Comparison Between Construction of Shorthand Outlines According to Theory and the Accuracy of Transcription." (Unpub. MS thesis, University of Washington, 1959.)
41. Mellinger, Morris. Basic Vocabulary for Written Business Office Communications. Chicago: Chicago State College, 1970.
42. Mellinger, Morris. "Has the Syllabic Intensity Yardstick Lost Its Magic?" Business Education World (November, 1964), 45, pp. 9-11.
43. Mellinger, Morris. "Let's Adopt the Yardstick of a Word Frequency Index." Business Education World (December, 1964), 45, pp. 13-15.
44. Meyer, Lois Irene. "A Test of the Validity of a Measure of Difficulty of Shorthand Dictation Materials." (Unpub. Ph.D dissertation, University of Minnesota, 1967.)
45. Mickelsen, Leonard P. "The Relationship Between Word Frequency and the Difficulty of Shorthand Dictation Materials." (Unpub. Ed.D dissertation, University of North Dakota, 1970.)
46. Minnick, Barbara Joy. "An Evaluation of Systematic Repetition of Brief Forms Through Specially Constructed Dictation Material for Gregg Shorthand, Diamond Jubilee Series." (Unpub. Ed.D dissertation, The University of Tennessee, 1967.)
47. Nickerson, Ingeborg Antonie. "A Measure of Difficulty of Shorthand Dictation Materials." (Unpub. Ed.D dissertation, Louisiana Tech University, 1977.)
48. Palmer, Rose. "A Comparison Between Two Groups of Shorthand Writers." (Unpub. Ph.D dissertation, New York University, 1965.)
49. Patrick, A. L. "An Error Analysis of Selected Brief Forms and Principles in Shorthand Notes of Beginning Gregg Diamond Jubilee Shorthand Students." (Unpub. Ed.D dissertation, The University of Tennessee, 1965.)



50. Perry, Devern J. "An Analytical Comparison of the Relative Word-Combination Frequencies of Business Correspondence With Phrase Frequencies of Selected Shorthand Textbooks." (Unpub. Ed.D dissertation, University of North Dakota, 1968.)
51. Peterson, Richard B. "An Investigation of the Validity and Reliability of a Formula for Determining Difficulty of Shorthand Dictation Materials." (Unpub. MA thesis, Mankato State College, 1964.)
52. Popham W. James and Sirotnik, Kenneth A. Educational Statistics--Use and Interpretation. 2nd Ed. New York: Harper & Row, 1973.
53. Pullis, Joe M. "A Test of the Validity of Syllabic Intensity and Percent High-Frequency Words as a Measure of Difficulty of Published Shorthand Dictation Materials." (Independent study, Louisiana Tech University, 1975.)
54. Pullis, Joe M. "A Test of the Validity of a Triple Control (Percent High-Frequency Words, Average Word Length, and Overall Syllable Intensity) As a Measure of Difficulty of Shorthand Dictation Materials." (Independent study, Louisiana Tech University, 1974.)
55. Pullis, Joe M. "Concerns of Shorthand Teachers: Part IV--Improving Evaluation of Shorthand Students." The Balance Sheet, LVIII (December 1976-January 1977), pp. 154-159.
56. Pullis, Joe M. "Concerns of Shorthand Teachers: Part II--Improving Teaching Methodology." The Balance Sheet, LVIII (October, 1976), pp. 64-66, 87.
57. Pullis, Joe M. "The Relationship Between Competency in Shorthand Accuracy and Achievement in Shorthand Dictation." (Unpub. Ed.D dissertation, North Texas State University, 1966.)
58. Pullis, Joe M. "Milking Some Sacred Cows of Shorthand." Century 21 Reporter (Fall, 1974), pp. 11, 15.
59. Pullis, Joe M. and Nickerson, Ingeborg. "Determining Dictation Difficulty: A Measurement Dilemma." Century 21 Reporter (Fall, 1975), pp. 11, 13-14.
60. Reese, Don and Smith, Ray. "A Model for Shorthand Dictation Material." Journal of Business Education, 50 (November, 1974), pp. 76-78.
61. Rowe, Clyde E. "The Writing of Infrequently Used Words in Shorthand." Contributions to Education, No. 869. New York: Columbia University Teacher's College, 1943.

62. Schramper, Frances Nies. "The Stroke Unit in Shorthand Measurement." Research Studies in Commercial Education, II (March, 1927), pp. 122-127.
63. Silverthorn, J. A. "The Basic Vocabulary of Written Business Communications." (Unpub. Ed.D dissertation, Indiana University, 1955.)
64. Summers, G. W. and Peters, W. S. Basic Statistics in Business and Economics. Belmont, Calif.: Wadsworth Publishing Co., Inc., 1973.
65. Tate, Donald J. and Hoskinson, Robert E., eds. Evaluation and Accountability in Business Education. E. Ray Smith, "Evaluating Instructional Objectives in the Stenographic Area." National Business Education Yearbook, No. 16. Reston, Va.: National Business Education Association, 1978. pp. 186-197.
66. Taylor, Wilson L. "'Cloze Procedure' A New Tool for Measuring Readability." Journalism Quarterly, 30 (Fall, 1953), pp. 415-433.
67. Thorndike, Edward L. The Teacher's Word Book. New York City: Teachers College, Columbia University, 1921.
68. Thorndike, E. L., and Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teachers College, Columbia University, 1944.
69. Tronsue, Mary. "A Study of Relationships of Selected Variables With Shorthand Transcription Speed." (Unpub. Ed.D dissertation, University of Maryland, 1968.)
70. Turse, Paul L. "Standard Word vs. Shorthand Stroke in Shorthand Dictation." Business Education World, XXIV (June, 1944), pp. 541-542.
71. Turse, Paul L. "The Validity of the Syllable Count." Journal of Business Education, XXIII (March, 1948), pp. 29-30, 32.
72. Uthe, Elaine F. "An Evaluation of the Difficulty of Shorthand Dictation Materials." (Unpub. Ph.D dissertation, University of Minnesota, 1966.)
73. Wedell, A. J. "The Relationship of Syllabic Intensity, Word Frequency, and Stroke Intensity to the Difficulty of Shorthand Dictation Material." (Unpub. Ph.D dissertation, The University of North Dakota, 1972.)
74. Wellman, Rowena. "An Examination of Certain Factors Involved in the Reporting and Transcribing of Stenographic Materials." (Published Ph.D dissertation, Columbia University, 1937.)

75. Wessman, A. L. "The Effect of Varying Time Lapses on Accuracy of Shorthand Transcription." (Unpub. MS thesis, University of Minnesota, 1956.)
76. West, Leonard J. "Principles and Procedures for Testing of Typewriting and Stenographic Proficiency." Business Education Forum, 29 (February, 1975), pp. 24-32.
77. Young, Israel. "An Analytical Study of the Effect of Choice Making on the Speed of Writing Symbol Combinations in Gregg and Pitman Shorthand." (Unpub. Ph.D dissertation, New York University, 1965.)
78. Zoubek, Charles E. Speed Dictation. New York: Gregg Publishing Division, McGraw-Hill Book Company, Inc., 1963.

APPENDIX A

FORM USED TO RECORD AND TABULATE THE NUMBER  
OF WORDS NEEDED AND/OR USED IN  
EACH CATEGORY CONTROLLED

Code:		GOAL	ACTUAL
1-100	Red		
101-500	Green		
501-1000	Blue		
1001-1500	Brown		
1501-3500	Pencil		
3501-5000	Pen		
Over 5000			
Total words			
Syllabic Inten.			
Brief Forms	Yel.		
B F Derivatives	C		
No. of Dif. words			

APPENDIX B

TESTS DEVELOPED FOR THIS STUDY

## EASY DIFFICULTY LEVEL STUDY TEST

Version 1

Dear Mrs. Hope

We appreciate your letter thanking us for the consideration and service that you received in our order department last week. It is always a pleasure to hear from happy customers and everyone in our group has asked me to send their thanks and appreciation to you. As you know, we have an experienced, professional staff that is interested in providing you with the finest quality merchandise you can purchase. Each piece is made and shipped with the greatest of care, and we guarantee that it will be received in satisfactory condition.

We are beginning a limited sale on several original special order table sets from a number of our new service companies. These convenient items have been going for more but will be reduced in price until all on hand are sold. They are very well made and cost no more than other products. The support is high enough to write on, and you can use them in any room. For example, one could be used for the telephone, or perhaps for letter writing or educational requirements. Two or more could be used with a desk or placed together between other items if you like.

This is the first time we have ever offered these items at such a substantial savings to customers. If further information would be of interest to you, why not visit our store now and perhaps select one for your home.

Sincerely yours

## EASY DIFFICULTY LEVEL STUDY TEST

Version 2

Dear Mr. West

Information about the construction of the new business building in the public part of the development along your street is certain to come from the Central Services Office about the end of the year. Once these arrangements have been made, we will have more budget to serve the city, including the freight and manufacturing business on either side of the road, better.

The new ten-inch pipe which you requested should also enable us to extend our service to other established companies that expect to add to their regular facilities. The excellent response and advice of your firm in working on this new plan is certainly appreciated. You have done much to keep our needs before the main office, and it has been a pleasure to work with your associates.

After I review the information in my records, I will make a copy and have it delivered to your office to be filed with the other papers. Perhaps it can help as a reference if a second section of meetings is held at a later time. As you know, that is a real possibility because of the fast growth in our area.

Thanks again for all your efforts in getting this important project accepted. We appreciate the service and interest of your company. If you hear that a decision has been reached, please let me or one of the members of the committee know immediately.

Sincerely yours



## AVERAGE DIFFICULTY LEVEL STUDY TEST

## Version 1

Dear Mrs. Moore

To satisfy public demand, we have added two important new services which we think will please our customers. Both are placed in convenient locations and are so easy to use that even a child can do them. Merely by following a few simple instructions and putting a coin in the box, you can have perfect copies of all your correspondence. Even the extreme letters are sharp and clear. Having access to this service will make time savings possible.

In addition to the copy machine, we have a new ink printer that can be used for duplicating reports, announcements, and other items in quantity. You need only put your original on the top, set the dial on the console for the number you want, and insert a quarter for every dozen copies. How many times have you wished you knew where to go to duplicate a club news bulletin or a party invitation? Now you can do it readily.

We think these new services will be very useful to several suppliers in the area also. Advertisements and circulars can be printed while they wait and for a fraction of the cost elsewhere. There is no purchase or maintenance expense on either machine.

We invite you to come by and try them out soon. Just raise the frame, put your masters on the plate, and drop in the appropriate amount. From beginning to end, it's a great duplicating value.

Cordially

## AVERAGE DIFFICULTY LEVEL STUDY TEST

Version 2

Dear Mr. Hall

Would you like to own a beautiful new home in the exclusive Western Hills region? If so, now may be the time. We have just listed a lake-side property there that is an excellent value for the money. It has good financing available, and it is ready for occupancy immediately. The home will be open for several hours tomorrow morning and we could probably see it then.

A brief newspaper ad is slated for publication this weekend. Because I feel certain an offer will be presented to the sellers within a day or two, I suggest you act quickly or the opportunity to buy will be gone.

I am enclosing a display folder that gives more information about the house and lot. You will notice that all appliances are included. Also, I understand that new carpets and drapes were recently installed throughout. The attached garage has a bank of deluxe storage cabinets and room for two cars.

The lot is approximately one acre and has several hundred feet of water front on the east. As you can see from the folder, there is a small private beach and boat house with service accommodations. The yard is completely landscaped and there are a number of beautiful trees on the property. If you wished, the area between the garden and the street could be enclosed by boundaries.

Do you have any further questions? If so, telephone me collect.

Sincerely

## HARD DIFFICULTY LEVEL STUDY TEST

## Version 1

Dear Mrs. Keys

An old saying tells that an ounce of prevention is worth a pound of cure. If that is accurate, a little work in your home now can prevent additional problems later on. With winter starting, you can expect the expense of heat and electric utilities to be substantially increased.

As you probably know, one area where partial heat loss occurs is under the floors. The enclosed circular shows how a cold draft can enter through openings in the window sills and most cracks under the doors. Your furnace must work harder and run longer to maintain a comfortable level of heat. The result is undue energy costs, more fuel used, and uneven temperatures.

The circular also shows how successful our regular repair strips are in stopping these drafts between the boards and screens. By attaching a block of two or more to every place that problems exist, you can eliminate each one easier than you would think. In addition, several strips joined together can be used for stopping big leaks and for repairing most damaged or broken places until replacements can be made.

These handy kits are great for lots of uses around your business, too, and are limited only by your needs. By following the simple pictures, you never need to worry about failure. For a chance to examine a giant sample, send your request on the block order coupon attached to the circular.

Yours truly

## HARD DIFFICULTY LEVEL STUDY TEST

## Version 2

Gentlemen

We will be staffing a big advertising organization in the immediate future and must find a trainee to handle the reception desk. Your expert staff has been sending us referrals for many years, and I am hoping you have prospects that meet our needs.

The person we employ must have neat appearance, shorthand skill, and good character. We also require filing skill and a nice telephone voice. The wages will depend on their background and the key hiring scale.

We'd like to promote this person to publication aide after a year. Therefore, we're reluctant to hire anyone with no experience or not trained in working with the publishing markets. We are equal opportunity employers and would try either a man or a woman. Age is no barrier to work with our group and neither is race or color. Word of this opening was made public yesterday morning in several bulk newspaper ads. A few brief replies have come in.

Among other fringe benefits, we have good health and accident plans and broad limits for sickness. Vacation time earned varies with years worked. Our employees can buy low-cost life insurance if they wish, and we have a good bonus check plan for those who retire after sixty. Chances for promotion depend on demand, experience, and related factors.

You can have applicants correspond with one of our shared testing centers if they would like an interview for the opening.

Cordially

APPENDIX C

TRANSCRIPT OF INTRODUCTORY MATERIAL ON  
THE SIX STUDY TEST TAPES

TRANSCRIPT OF INTRODUCTORY MATERIAL ON  
THE SIX STUDY TEST TAPES

This is a shorthand test tape at 80 words per minute for three minutes. Make a note of the following four items which should be typed in the upper right hand corner of your transcription paper.

1. Your name
2. Your school
3. Today's date
4. The code (A, B, or C) 80-(1 or 2)

To review, that's your name, your school, today's date, and the code (A, B, or C) 80-(1 or 2).

Please use a one-inch margin with double spacing for your transcript.

You will now be given a brief warmup and then the take. Record as much of the dictation as you can. You should continue writing even if you feel you have missed so much you cannot pass. It is also important that you transcribe every word that you can. The more you can transcribe, the higher your score will be.

Remember, you should record and transcribe everything you possibly can, regardless of whether your paper is passing or not.

Now make sure you are comfortable and relaxed. Here is the warmup.

(Warmup at 100 words per minute for one minute)

Get ready for the three-minute take now. (Pause) Ready.

(Give take at 80 words per minute for three minutes)

APPENDIX D

GUIDELINES FOR TEACHERS

## GUIDELINES

The following guidelines may be helpful for those administering the tests for this study.

1. Normal classroom atmosphere should be maintained as closely as possible. If students ask, they may be told that this is a special series of takes. Any other responses to questions should be very general. To avoid contamination of the results, it is preferred that students know very little about the design or purpose of the study.
2. Materials should not be previewed and students should not be permitted to practice the testing materials in any way.
3. Students are instructed on the tape to use one-inch margins and double spacing for the transcript. They are also told to put the following information in the upper right hand corner of their paper: Name, School, Current date, and the code for the particular test they are taking. The above data is for purposes of recordkeeping and accuracy in collecting scores. No students or schools will be specifically identified by name or other means in the results of the study.
4. Each of the two preliminary takes contains instructions to the students, a brief warmup, and two 3-minute dictations. One preliminary take is at 80 words per minute and the other is at 100 words per minute. Each of the six regular test tapes contains instructions to the students, a brief warmup, and one 3-minute dictation at 80 words per minute.
5. The purposes of the preliminary tests are to establish skill levels and to show that students are approximately equal in ability. Results of only the better of each of the two preliminary takes at each speed will be used to establish each student's skill level.
6. Students should record and attempt to transcribe both of the preliminary takes at 80 words per minute on the first day of testing. They should follow the same procedure for both of the 100 takes on the second. They are instructed on the tape to transcribe the better of the two first and then spend the rest of the class time working on the other.
7. The six regular takes should be given during the two weeks following the preliminary takes in the order indicated at the rate of three takes per week. They should be given on consecutive days, avoiding Monday wherever possible.
8. Only the shorthand errors will be counted for this study. This includes words that are omitted, inserted, mistranscribed, or otherwise changed from the verbatim dictation. Punctuation, capitalization, typographical, spelling, and other grammatical errors will not be counted.
9. Please encourage students to transcribe every word that they can even though the take may not be passing. They should not leave early or



quit before the class period is over. All papers will be evaluated on total words so every correct word transcribed is helpful.

10. In order to get maximum effort, students should feel that these tests will be graded and will count for points or grades.
11. The transcription need not be timed for purposes of this study, and students may correct errors if they wish. However, the transcription should be completed during the same class period in which it is dictated.
12. Because only transcripts for students who have taken all eight tests can be used, you may administer makeup tests during the same week if it is convenient to do so.
13. Student's shorthand notes should be collected and stapled to the back of the transcript at the end of each day. The transcript, notes, and test tapes should be returned to the researcher in the mailing envelope provided as soon as possible.
14. All correcting and grading will be done by the researcher. If you wish to have scores for your students sent to you, please indicate that fact in a note with the transcripts when they are submitted. Scores will be returned to you as soon as they are available.
15. Should you have any questions or problems, please feel free to call me collect at (405) 377-5617. My home address is D3-6 Brumley Apartments, Stillwater, OK 74074.

APPENDIX E

TRANSCRIPT OF INTRODUCTORY MATERIAL ON  
THE PRELIMINARY TEST TAPES

TRANSCRIPT OF INTRODUCTORY MATERIAL ON  
THE PRELIMINARY TEST TAPES

This is a shorthand test tape at (80 or 100) words per minute for three minutes. Make a note of the following four items which should be typed in the upper right hand corner of your transcription paper.

1. Your name
2. Your school
3. Today's date
4. The code P-(80 or 100)

To review, that's your name, your school, today's date, and the code P-(80 or 100).

Please use a one-inch margin with double spacing for your transcript.

You will now be given a brief warmup and then two takes at (80 or 100) words per minute. Record as much of the dictation as you can. You should continue writing even though you may feel you have missed so much you cannot pass. Transcribe the take on which you feel you did best first, and then transcribe as much as you can of the other one in the time you have left.

It is important that you transcribe every word that you can. The more you can transcribe, the higher your score will be. Remember, you should record and transcribe everything you possibly can regardless of whether your paper is passing or not.

Now, make sure you are comfortable and relaxed. Here is the warmup.

(One-minute warmup at 100 or 120)

Get ready for the first three-minute take now. (Pause) Ready.

(Give take at 80 or 100)

(When take is completed, pause for approximately 1 minute)

Here is the second three-minute take at (80 or 100) words per minute. Ready.

(Give second take at 80 or 100)

APPENDIX F

FORM USED TO RECORD ERROR SCORES



APPENDIX G

SUMMARY OF RAW SCORES FOR ALL STUDENTS  
ON SIX STUDY TESTS

TABLE XIII  
 SUMMARY OF RAW SCORES FOR ALL STUDENTS  
 ON SIX STUDY TESTS

Student Number	Class	Easy		Average		Hard	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
1	1	17	35	44	45	97	69
2	1	7	11	22	9	70	62
3	1	15	24	38	24	113	50
4	1	2	5	13	6	38	45
5	1	13	2	21	12	60	57
6	1	16	13	18	15	36	50
7	1	2	2	1	3	14	17
8	1	17	42	38	40	67	76
9	1	13	13	21	45	40	41
10	1	2	4	36	24	21	31
11	1	23	46	69	49	78	69
12	1	12	2	7	19	20	28
13	1	8	1	3	7	7	15
14	2	4	2	9	14	53	74
15	2	1	3	2	4	10	12
16	2	3	2	5	14	17	21
17	2	2	5	6	5	16	22
18	2	0	2	5	5	3	2
19	2	2	0	0	4	5	9
20	2	1	5	4	16	32	20
21	2	0	2	5	4	12	6
22	2	2	4	4	1	10	15
23	2	16	13	23	30	73	50
24	2	10	23	35	16	62	67
25	2	7	1	5	4	17	15
26	2	8	1	22	29	42	12
27	2	1	0	6	5	19	8
28	2	8	0	11	10	21	43
29	2	0	4	6	4	16	5
30	2	2	4	3	3	12	16
31	2	2	1	11	5	31	14
32	3	21	25	30	27	74	49
33	3	11	6	28	28	61	55
34	3	4	5	22	22	79	65
35	3	2	0	6	4	15	7
36	3	2	1	6	7	32	21
37	3	8	26	38	12	42	43
38	3	9	7	25	16	36	21
39	3	6	6	32	33	82	65
40	3	4	6	13	32	31	28

TABLE XIII (Continued)

Student Number	Class	Easy		Average		Hard	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
41	3	12	1	23	17	15	37
42	3	5	2	1	5	3	0
43	3	17	6	23	11	56	57
44	3	3	16	8	4	21	12
45	3	10	12	40	13	39	37
46	3	11	4	32	13	46	61
47	3	5	24	8	14	75	54
48	3	33	10	31	8	42	26
49	3	9	1	10	10	37	32
50	3	8	26	19	10	44	23
51	3	17	9	31	51	98	75
52	3	14	15	17	30	53	52
53	4	7	5	4	7	21	20
54	4	9	3	23	17	41	46
55	4	20	6	17	22	59	34
56	4	7	4	39	51	84	86
57	4	14	9	33	40	55	65
58	4	1	0	3	2	12	3
59	4	2	5	5	14	27	18
60	4	0	1	2	14	11	17
61	4	0	3	7	3	23	12
62	5	4	11	21	19	42	35
63	5	5	2	13	2	13	32
64	5	9	22	29	27	57	108
65	5	23	41	48	45	112	47
66	5	22	10	38	24	64	41
67	5	4	16	7	4	14	21
68	5	1	27	9	17	41	41
69	5	2	2	18	6	29	4
70	5	1	1	1	2	4	10
71	5	4	0	1	3	7	6
72	5	17	11	17	16	22	39
73	5	1	4	3	5	22	16
74	5	22	32	38	43	93	66
75	6	3	1	23	2	45	57
76	6	9	6	18	22	62	95
77	6	1	5	11	16	51	70
78	6	25	10	35	17	96	104
79	6	2	3	2	10	33	52
80	6	15	27	54	39	83	76
81	6	7	15	12	38	40	44
82	6	20	22	53	60	115	109
83	6	5	9	8	8	18	26
84	6	26	39	78	60	123	122
85	6	15	24	40	45	87	77



TABLE XIII (Continued)

Student Number	Class	Easy		Average		Hard	
		Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
86	6	0	4	21	17	24	55
87	6	1	1	4	14	24	41
88	6	21	11	20	14	51	66
89	6	1	4	4	7	28	27
90	6	14	7	5	3	79	50
91	6	3	2	11	18	34	23
92	6	37	14	11	57	91	102
93	7	8	12	40	47	87	66
94	7	14	7	16	12	27	21
95	7	21	6	35	31	31	52
96	7	3	13	13	22	43	49
97	7	14	9	21	19	52	63
98	7	12	9	45	34	66	75
99	7	14	40	72	67	115	83
100	7	17	8	38	43	66	68
101	8	12	2	7	13	36	57
102	8	7	9	8	18	159	139
103	8	7	9	16	10	41	26
104	8	10	8	22	19	68	63
105	8	8	5	28	14	30	56
106	8	13	15	42	33	111	66

APPENDIX H

SUMMARY OF RAW MEAN TRANSCRIPTION ERROR SCORES  
FOR EIGHT CLASSES

TABLE XIV  
 SUMMARY OF RAW MEAN TRANSCRIPTION ERROR SCORES  
 FOR EIGHT CLASSES

Class	Easy		Average		Hard	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
1	11.31	15.38	25.46	22.92	50.85	46.92
2	3.83	4.00	9.00	9.61	25.06	22.83
3	10.05	9.91	21.10	17.48	46.71	39.05
4	7.11	4.00	14.78	18.89	37.00	33.44
5	8.85	13.78	18.69	16.38	40.00	35.85
6	11.39	11.33	22.78	24.83	60.22	66.44
7	12.88	13.00	35.00	34.38	60.88	59.63
8	9.50	8.00	20.50	17.83	74.17	66.83

APPENDIX I

COMPUTATION OF Z SCORES FOR THREE  
DIFFICULTY LEVEL COMPARISONS

TABLE XV  
 COMPUTATION OF Z SCORES FOR THREE  
 DIFFICULTY LEVEL COMPARISONS

Difficulty Level	Square A	Square B	Square C	Square D
Average - Easy	$\frac{9.517}{1.473}$	$\frac{6.552}{1.473}$	$\frac{9.318}{1.583}$	$\frac{9.961}{1.583}$
Hard - Average	$\frac{22.83}{2.074}$	$\frac{19.64}{2.074}$	$\frac{30.28}{2.147}$	$\frac{29.96}{2.147}$
Hard - Easy	$\frac{32.35}{1.97}$	$\frac{26.19}{1.97}$	$\frac{39.59}{2.065}$	$\frac{39.92}{2.065}$

## VITA<sup>2</sup>

Annette Johnson Thomason

Candidate for the Degree of

Doctor of Education

Thesis: THE EFFECT OF WORD FREQUENCY ON COPY DIFFICULTY FOR SHORTHAND TESTING MATERIALS

Major Field: Business Education

### Biographical:

Personal Data: Born in Provo, Utah, October 1, 1937, the daughter of Marion P. and Verl J. Johnson; Mother of two sons: Allen and Mark.

Education: Graduated from Provo High School in Provo, Utah, May 1955; received a Secretarial Diploma from LDS Business College, Salt Lake City, Utah, June, 1956; received the Bachelor of Science degree magna cum laude from Brigham Young University, Provo, Utah, in May, 1969, with a major in Business Education; received the Master of Science degree from Brigham Young University, Provo, Utah, in August, 1971, with a major in Business Education and Related Fields; certified as a Certified Professional Secretary in May, 1972; completed the requirements for the Doctor of Education degree at Oklahoma State University in May, 1979.

Professional Experience: Eight years professional secretarial experience in positions ranging from steno to administrative assistant prior to 1965; business teacher at American Fork High School, American Fork, Utah, 1969-1971; business instructor at Utah Technical College at Provo, Provo, Utah, 1971-1977; graduate teaching associate in Department of Administrative Services and Business Education at Oklahoma State University, Stillwater, Oklahoma, 1977-1978.

Professional Memberships: National Business Education Association, Western Business Education Association, National Education Association, American Vocational Association, Utah Business Education Association, Utah Education Association, Utah Vocational Association, National Secretaries Association, Delta Pi Epsilon, Phi Kappa Phi, Beta Gamma Sigma, Phi Delta Kappa.