A COMPARATIVE STUDY OF THE COMPETENCE OF GROUPS OF INTERNATIONAL AND NATIVE STUDENTS IN ASPECTS OF LANGUAGE THAT HOLD SPECIAL RELEVANCE TO SPEECH

DISSERTATION

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

By

ROBERT LEONARD MULDER, B.A., M.A.
The Ohio State University

1953

Approved by:

[Signature]
Adviser
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Organization of the dissertation</td>
<td>15</td>
</tr>
<tr>
<td>II. REVIEW OF THE LITERATURE</td>
<td>17</td>
</tr>
<tr>
<td>1. Studies which treat the testing of comprehension of English as a foreign language</td>
<td>17</td>
</tr>
<tr>
<td>2. Studies which treat the relationship between speech-sound discrimination and articulatory proficiency</td>
<td>27</td>
</tr>
<tr>
<td>3. Studies which deal with measures of intelligibility</td>
<td>34</td>
</tr>
<tr>
<td>4. Studies related to indirect methods of assessing articulation proficiency</td>
<td>40</td>
</tr>
<tr>
<td>5. Studies which treat prediction of letters</td>
<td>42</td>
</tr>
<tr>
<td>Summary of the chapter</td>
<td>45</td>
</tr>
<tr>
<td>III. METHODS AND PROCEDURES FOR SELECTION AND PREPARATION OF THE MATERIALS, SELECTION OF SUBJECTS, AND TESTING OF SUBJECTS</td>
<td>46</td>
</tr>
<tr>
<td>Part I. The selection and preparation of materials</td>
<td>47</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>PAGE</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Comprehension of English: aural comprehension</td>
<td>47</td>
</tr>
<tr>
<td>Speech-sound discrimination</td>
<td>50</td>
</tr>
<tr>
<td>Intelligibility</td>
<td>52</td>
</tr>
<tr>
<td>Comprehension of English: letter prediction</td>
<td>55</td>
</tr>
<tr>
<td>Controlled speech patterns</td>
<td>58</td>
</tr>
<tr>
<td>Stimulus material for controlled speech patterns</td>
<td>59</td>
</tr>
<tr>
<td>Speakers</td>
<td>61</td>
</tr>
<tr>
<td>Recording the phrases</td>
<td>62</td>
</tr>
<tr>
<td>Construction of Tape A and Tape B</td>
<td>63</td>
</tr>
<tr>
<td>Assignment of scale values to the Tapes A and B</td>
<td>64</td>
</tr>
<tr>
<td>The reliability of Tape I (nine-second passages)</td>
<td>69</td>
</tr>
<tr>
<td>The validity of Tape I (nine-second passages)</td>
<td>71</td>
</tr>
<tr>
<td>The construction of Tape III (pairs)</td>
<td>72</td>
</tr>
<tr>
<td>Summary of Part I</td>
<td>74</td>
</tr>
<tr>
<td>Part II. Selection of the experimental subjects</td>
<td>75</td>
</tr>
<tr>
<td>Section A. Selection of sixty native and non-native speakers of English</td>
<td>75</td>
</tr>
</tbody>
</table>
### CHAPTER

<table>
<thead>
<tr>
<th>Section B. Selection of thirty-six non-native speakers of English</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>78</td>
</tr>
</tbody>
</table>

| Part III. The procedures followed in testing                | 82   |
|                                                               | 83   |
| Summary of Chapter III                                       | 83   |

| IV. RESULTS AND DISCUSSION                                   | 85   |
|                                                               |      |
| Part I. Results based on scores obtained for sixty subjects | 85   |
|                                                               | 85   |
| The computation of the scores for the six tests              | 85   |
|                                                               | 85   |
| Aural comprehension                                          | 85   |
|                                                               | 85   |
| Speech-sound discrimination                                  | 85   |
|                                                               | 86   |
| Tape I (nine-second passages)                                | 86   |
|                                                               | 86   |
| Tape III (pairs)                                             | 86   |
|                                                               | 87   |
| Intelligibility                                              | 87   |
|                                                               | 87   |
| Letter prediction                                            | 87   |
|                                                               | 88   |
| Graphic presentation of the scores                           | 99   |
| Hypothesis I                                                 | 99   |
|                                                               | 107  |
| Hypothesis II                                                | 117  |
|                                                               |      |
| Graphic presentation of the scores for the test of letter-prediction ability | 107 |
| Hypothesis III                                               | 117  |

| Part II. Results based on scores obtained for thirty-six international students | 122 |
|                                                                                   | 122  |
Hypothesis IV ............................... 122

The reliability of Tape I (nine-second passages for international students) .................. 128

V. CONCLUSIONS ............................ 130

Selection and preparation of the materials for presentation to the subjects ..... 131

Selection of the experimental subjects ........................................... 134

Hypothesis I ..................................... 134

Hypothesis II .................................... 136

Hypothesis III ................................... 136

Hypothesis IV .................................... 138

1. Comprehension of English, as measured by Lado's Test of Aural Comprehension, is related to speech intelligibility. 141

2. As far as the native speakers of Spanish and Japanese selected for this study are concerned, intelligibility is not affected by language background . . . . 142

3. Tape I (nine-second passages) is a valid and reliable indirect test of intelligibility . . . . . . . . . . . . . 143
4. Shannon's statement can be modified to read: "Anyone knowing a language possesses an enormous knowledge of the statistics of a language and this knowledge varies with his comprehension of the language."
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Scale Values, Mean Scale Values, and Mean Q-Values Assigned to Tape A by 100 College Students (Group A)</td>
<td>67</td>
</tr>
<tr>
<td>II. Assigned Scale Values and Mean Scale Values for Tape I (nine-second passages) for Three Groups of Judges</td>
<td>68</td>
</tr>
<tr>
<td>III. Summary of the Statistical Methods Employed in Analyzing the Obtained Data</td>
<td>89</td>
</tr>
<tr>
<td>IV. t Ratios for the Difference Between Obtained Means ($M_o$) and Guessed Means ($M_g$) for Three Groups of Subjects for Six Tests</td>
<td>98</td>
</tr>
<tr>
<td>VI. t Ratios for Tests of Significant Differences among the Means of Three Groups for Six Measures</td>
<td>101</td>
</tr>
<tr>
<td>VII. The Means, Standard Deviations, and Standard Error of the Means Obtained from the Data for Letter Prediction Based on the Number of Letters Predicted Correctly on the First Trial for Ninety Letters</td>
<td>103</td>
</tr>
</tbody>
</table>
TABLE

VIII. Summary of Analysis of Variance of Scores Based on the Number of Letters Predicted Correctly on the First Trial by Native Speakers of Spanish, Japanese, and English ........................................ 104

IX. The Means, Standard Deviations, and Standard Error of the Means Obtained from the Data for Letter Prediction Based on the Number of Letters Predicted Correctly after Ten Previous Trials for Ninety Letters ........ 105

X. Summary of Analysis of Variance of Scores Based on the Number of Letters Predicted Correctly after Ten Previous Trials by Native Speakers of Spanish, Japanese, and English ................................... 106

XI. Product-moment Correlations Based on the Scores for Sixty Native Speakers of Spanish, Japanese, and English .......................... 119

XII. Product-moment Correlations Based on the Scores for Forty Japanese- and Spanish-speaking Subjects ............................. 120
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Plot of Cumulative Scores Obtained for Twenty Native Speakers of Spanish, Twenty Native Speakers of Japanese, and Twenty Native Speakers of English for the Test of Aural Comprehension</td>
<td>90</td>
</tr>
<tr>
<td>2. Plot of Cumulative Scores Obtained for Twenty Native Speakers of Spanish, Twenty Native Speakers of Japanese, and Twenty Native Speakers of English for the Test of Sound Discrimination</td>
<td>92</td>
</tr>
<tr>
<td>3. Plot of Cumulative Scores Obtained for Twenty Native Speakers of Spanish, Twenty Native Speakers of Japanese, and Twenty Native Speakers of English for Tape I (nine-second passages)</td>
<td>93</td>
</tr>
<tr>
<td>4. Plot of Cumulative Scores Obtained for Twenty Native Speakers of Spanish, Twenty Native Speakers of Japanese, and Twenty Native Speakers of English for Tape III (pairs)</td>
<td>94</td>
</tr>
<tr>
<td>5. Plot of Cumulative Scores Obtained for Twenty Native Speakers of Spanish, Twenty Native Speakers of Japanese, and Twenty Native Speakers of English for the Waco Multiple-Choice Intelligibility Test</td>
<td>96</td>
</tr>
</tbody>
</table>
6. Plot of Cumulative Scores Obtained for Twenty Native Speakers of Spanish, Twenty Native Speakers of Japanese, and Twenty Native Speakers of English for the Test of Letter Prediction

7. Mean Number of Trials for the Prediction of the Letters "in and around t"

8. Mean Number of Trials for the Prediction of the Letters "on the east fri"

9. Mean Number of Trials for the Prediction of the Letters "here at a meeti"

10. Mean Number of Trials for the Prediction of the Letters "who have been d"

11. Mean Number of Trials for the Prediction of the Letters "was taken becau"

12. Mean Number of Trials for the Prediction of the Letters "to decide wheth"

13. Mean Number of Trials Per Sequence of Three Letters (phrases pooled)

14. Mean Number of Ninety Letters Predicted Correctly on the First Trial and After Ten Previous Trials for Three Groups of Experimental Subjects
15. Distribution of Scores for Thirty-six Non-native Speakers of English for Test of Aural Comprehension

16. Distribution of Scores for Thirty-six Non-native Speakers of English for Tape I (nine-second passages)

17. Distribution of Scores for Thirty-six Non-native Speakers of English for Waco Write-down Test of Intelligibility
CHAPTER I

INTRODUCTION

This study was designed to test several hypotheses concerning English as a foreign language. The hypotheses pertain to the relationships and differences among measures of comprehension of English, speech intelligibility, judgments of controlled speech patterns, and speech-sound discrimination ability. The results of this study are based on the performance of native and non-native speakers of English.

Since World War II there has been a large increase in the number of international students attending colleges and universities in the United States. The presence of these students on the American campus is attributable in part to the influx of displaced persons, in part to provisions of the Fulbright Act, and in part to the increasing relative importance of the academic resources of the United States. Many of the students find their way to a Department of Speech or Department of English because of their inability to communicate adequately in English. Thus, colleges which formerly had only an occasional international student are now developing special courses or programs in communication for international students.

During the Autumn Quarter, 1952, there were 258 students representing fifty-one nations on The Ohio State
University campus. These numbers do not include ninety-

1 "Directory of International Students, Autumn
Quarter, 1952," The Committee on International Students, 104
Administration Building, The Ohio State University, Columbus,
10, Ohio.

one post-doctoral Fellows, displaced persons, Puerto Ricans,
and Hawaiians. During this same quarter at least seventy-
five of the students were served in some professional capaci­
ty by staff members of the Department of Speech of The Ohio
State University.

International students in these numbers create special
problems usually not covered in the academic or service
programs of the Department of Speech or Department of
English. Many of the problems center around standards of
proficiency in respect to the ability to communicate
effectively.

When an international student attempts to learn to
speak a foreign language, he not only must be able to un­
derstand the new language as spoken vernacularly, but also
must speak the language so that he can be understood by the
native speaker.

Standardized techniques have been developed to measure
the foreign student's comprehension of written English. 2,3

2 Inter-American Tests. Princeton, New Jersey (P.O.
Box 592): Educational Testing Service.
But since the international student must understand language when spoken, tests have been developed to assess the student's ability to comprehend the language when spoken by a native speaker.  

Because the student should learn to speak the language as spoken by the native, he should learn the intricacies of pronunciation of the language. Although he should learn about variations in inflections and rhythm, knowledge relative to the similarities and differences of the sound system of his vernacular and the second language will be important to him. A few of the sounds of most languages are the same and require no new learning.  

Trubetzkoy suggested that designers of international languages should limit themselves to fourteen phonemes, five vowels and nine consonants, because these sounds, or recognizable approximations of them, can be found in all languages of major importance.  

language are not the same as those in the first. These sounds are the ones that cause difficulty when pronunciation is considered.

In learning to speak a second language, it is important that the speaker be aware of the acoustical patterns of the sounds of the two languages so that he can make the appropriate motor adjustments for speech. The ability to differentiate between the sounds of a language is called "sound discrimination ability."

The importance of sound discrimination is considered an important part of language proficiency even for native speakers of a language. Travis and Rasmus,7 Carrell,8 Hall,9 Hansen,10 Templin,11 Reid,12 Anderson,13 and others

have studied the sound discrimination ability of normal speakers of English and speakers of English with defective articulation. Although the results of these studies are inconsistent,\(^{14}\) the investigations were conducted because of the assumption that sound discrimination is essential to correct speech.

As far as could be determined, studies similar to those mentioned above have not been conducted relative to the articulatory ability of the non-native speaker of English and his ability to discriminate between sounds of English. Part of this investigation will be concerned with the ability of the non-native speaker of English to discriminate between the sounds of English and the relationship of this ability to intelligibility.
In recent years Agard and Dunkel and Lado have described tests of pronunciation. Both of these tests were designed for individual testing. Lado provides for the testing of specific sounds in context while Agard and Dunkel provide for an opinion of "Pronunciation and Enunciation" based on a three-point scale. Both of these tests rely on a single judge to determine the pronunciation ability of the international students. This is a questionable practice in terms of reliability.

The evaluation, for classroom purposes, of the speech of international students could be accomplished by a panel of judges. This type of evaluation is usually prohibitive, however, for the classroom teacher because of the problems related to assembling a panel of judges. A more practicable method would be that of assessing the speech of students in groups.

Lado has suggested indirect methods for testing the pronunciation of foreign students. In the section on...
pronunciation of the English Language Test for Foreign Students, the items consist of pairs of words. The student decides whether the words rhyme. In the remaining items, the students check the syllable that receives the most stress in words. The reliability of this kind of approach to pronunciation has been verified only in terms of the responses of one person. Lado reports that he had one student take the test. Before the paper was scored, Lado had the student read the test aloud while he scored his pronunciation. He found that the correlation for the pronunciation test score and the score obtained for reading was .92 (split-half corrected by the Spearman-Brown prophecy formula).

Black's Pronunciation Judgment Test is a recorded test and is divided into four parts. Part 1 contains items presented in a paired manner, with the word said well in one instance and with a common mispronunciation in the other. Part 2 consists of a list of words. Each word is said once.
The student indicates whether he considers the pronunciation of the word to be **good** or **poor**. The items in Part 3 are spoken in context. Again the student indicates whether he considers the pronunciation of the word to be **good** or **poor**. In Part 4, the student notes whether the accent of words falls on the first or second syllable.

The **Pronunciation Judgment Test** was found to be reasonably valid as an indirect test of pronunciation. This was determined by recording the responses of one hundred subjects as they read fifty-two of the test words from the recorded test. These responses were judged by four teachers of speech. The correlation for the error scores in reading and the error scores for the recorded test was, $r = .72$.

Pollikoff\(^{21}\) employed the **Pronunciation Judgment Test**

---


---

This test is too difficult for the average foreign student but is suggestive of the possibilities for indirect methods of testing pronunciation.
Blattner\textsuperscript{22} developed a series of group tests of pronunciation. The tests consisted of phonographed items of correct and incorrect pronunciation of single word productions. She did not, however, establish whether there was a relationship between judgments on the test and the way the subjects talked.

It will be noted that the above studies assume one or both of two things: (1) that what the person thinks he says is related to the way he speaks, and (2) that his preference for speech patterns is related to the way he speaks. These studies, however, are either too difficult or are not reliable enough for use with international students. Part of the present study is concerned with the construction of a reasonably valid and reliable indirect test of pronunciation suitable for use with international students.

Another approach to the measurement of competence in language is suggested by the comments made by Shannon.

\textit{... Anyone speaking a language possesses, implicitly, an enormous knowledge of the statistics of the language. Familiarity with the words, idioms, cliches and grammar enables him to fill in missing or incorrect letters in proof-reading, or to complete an unfinished phrase in conversation.}\textsuperscript{23}
In the same article in which the above statement was made, Shannon described an experimental demonstration of the extent to which English is predictable. A short passage unfamiliar to the person who was to do the predicting was selected. He was asked to guess the first letter in the passage. If the guess was correct he was so informed, and proceeded to guess the second letter. If the guess was not correct, he was told the correct letter and proceeded to guess the second. In one such experiment Shannon found that eighty-nine out of one hundred and twenty-nine letters (or 69 per cent) were guessed correctly on the first attempt.

He found that "The errors occur most frequently at the beginning of words and syllables where the line of thought had more possibility of branching out."

An extension of the above experiment required the subject to guess the letter as above. If he was wrong, he was informed and requested to continue guessing until he found the correct letter. Out of one hundred and two symbols, the subject in this experiment guessed right seventy-nine times on the first guess (or 77 per cent).

---

24 Ibid., p. 6.
25 Ibid., p. 7.
The sort of prediction developed by Shannon is amenable to statistical treatment because of the restrictions inherent in languages. Language may be considered as a Markoff chain composed of an ordered system of symbols.

A system which produces a sequence of symbols ... according to certain possibilities is called a stochastic process and the special case of a stochastic process in which the probabilities depend on the previous events is called a Markoff process or Markoff chain.26


Symbols in a Markoff chain may be words, sounds, letters of the alphabet, or other units. In any event, these symbols are ordered, i.e., they do not ordinarily occur in random sequences.

A few examples from the English language may help to clarify what is meant by the restrictions inherent in language. Word order is important in English. Hence, certain types of words have a very low probability of occurring after certain other words. For example, there is little likelihood that a verb directly would follow an article. On the other hand, it is very probable that a verb would precede and/or follow a noun.
Dewey\textsuperscript{27} found that in the sound system of English, vowels and diphthongs occur 38 per cent of the time, whereas, consonants occur 62 per cent of the time. Even among the consonants the phoneme $[n]$ occurs 7.24 per cent of the time, but the $[q]$ occurs only 0.05 per cent of the time. According to Newman,\textsuperscript{28} consonants and vowels occur in sequential patterns. He selected passages from the Bible in eleven languages. Coding the consonants and vowels in these passages in a series of binary items, he demonstrated marked periodicity in the alternation of vowels and consonants.

If there were no restrictions on a language,\textsuperscript{29} it would be expected that a person would require about half as many trials as there are letters in the language to predict a given letter. In other words, if the letters were from the English orthographic alphabet, he would have twenty-six letters to choose from and by chance alone he would be expected to find the correct letter after exhausting about one


\textsuperscript{29} It would then no longer be a language according to the definition above.
half of the possibilities. No change would occur in the number of trials necessary to predict the next letter as the selection would still be a matter of chance since there are no restrictions on the language. This is not the case according to Shannon, who found that 69 per cent of 129 letters were predicted correctly on the first attempt in one instance and that 77 per cent of 102 letters were guessed correctly on the first attempt in another instance.

If it is true, that anyone speaking a language possesses an enormous knowledge of the statistics of a language, then it would follow that the more sophisticated the person is with respect to a given language, the more of the statistics of the language he knows. American, therefore, should be able to predict events in a Markoff chain (the English language) more readily than foreigners who know English imperfectly. In other words, the better a person comprehends a language, the better he should be able to predict the letters of the alphabet of the language. A modification of Shannon's technique and its application to the testing of the assumption that Americans can predict the letters of the alphabet of English better than foreigners will be described in later chapters.
As stated at the beginning of this chapter, this study was designed to test several hypotheses related to English as a foreign language. It has been shown that no studies have been conducted relative to the sound discrimination ability of native and non-native speakers of English, even though sound discrimination ability is considered to be an important factor in the correct production of speech. Part of this study, therefore, will be concerned with testing the assumption that there is a difference between non-native and native speakers in the ability to discriminate between the sounds of English.

As suggested earlier, a group test of the pronunciation ability of students would be desirable for classroom purposes. An approach to the construction of a test of this type is that of noting what the person thinks he says, or what his preference is for speech patterns, and to relate these responses to the way he talks. Part of this study relates to the problem of the construction of a test of this sort.

Finally, it was suggested that Americans should be able to predict the letters of the alphabet of connected passages of English more readily than non-native speakers of English. Part of this study will be concerned with testing this assumption.
In order to test these assumptions the following null hypotheses were formulated:

1. There is no difference between native and non-native speakers of English (a) in the ability to predict the letters of English, (b) in the ability to comprehend English, (c) in the ability to differentiate between the sounds of English, (d) in intelligibility when speaking English, and (e) in the ability to judge controlled patterns of speech.

2. There is no difference between native speakers of Spanish and Japanese in the variables that are listed in 1 immediately above.

3. There is no relationship among the variables that are listed in 1 immediately above.

4. Judgments of controlled speech patterns are not reliable or valid estimates of intelligibility.

ORGANIZATION OF THE DISSERTATION

This dissertation reports a study which is designed to investigate several hypotheses related to the learning of English as a foreign language. It is divided into five chapters with references, appendices and an autobiography added. Chapter I summarizes the hypotheses under study and is concluded by this brief outline of the dissertation. Chapter II contains a brief review of the studies related to this dissertation. Chapter III outlines the selection and preparation of the materials for the study, the selection of
subjects, and the procedures followed in testing the sub-
jects. Chapter IV presents the results of the study.

Appendix A contains copies of the examination booklet
and answer sheet for the test of aural comprehension.
Appendix B contains a copy of the pairs of stimulus syllables,
the directions, and a sample answer sheet for the test of
sound discrimination. Appendix C contains a copy of the
lists of words read by the subjects for the Waco twenty-
four word multiple-choice test of intelligibility, Form B,
and a copy of the answer sheet for the listener-judges.
Appendix D contains a copy of the answer sheet for testing
letter-prediction ability. Appendix E contains a list of
270 phrases which were presented to 100 college students for
judgments of pronunciation adequacy. Appendix F contains a
copy of the word lists for the Waco twenty-four word write-
down test of intelligibility and a copy of the answer sheet
for the judges. Appendix G contains the raw scores for the
sixty native speakers of English, Spanish, and Japanese
for six tests. Appendix H contains the raw scores for
thirty-six non-native speakers of English.
CHAPTER II

REVIEW OF THE LITERATURE

This study was designed to investigate several hypotheses related to English as a foreign language. The studies in the literature relevant to the investigation herein described may be divided into five categories:

1. Studies that treat the testing of comprehension of English as a foreign language.
2. Studies that treat the relationship between speech-sound discrimination and articulatory proficiency.
3. Studies that treat measures of intelligibility.
4. Studies that treat indirect methods of assessing articulatory proficiency.
5. Studies that treat the prediction of the letters of English.

1. STUDIES THAT TREAT THE TESTING OF COMPREHENSION OF ENGLISH AS A FOREIGN LANGUAGE.

The Educational Testing Service, Princeton, New Jersey, has published two sets of tests relative to testing competence in English as a foreign language. The first of these tests is entitled English Examination for Foreign Students.¹ This test is designed to give educational

¹ English Examination for Foreign Students, Princeton, New Jersey, (P.O. Box 392). Educational Testing Service.
institutions information about the knowledge of English which foreign students have before admittance. The Educational Testing Service provides the materials to the admitting institution for a nominal charge and scores the objective parts of the test. The examination consists of five sections: (A) Reading Comprehension, (B) Aural Comprehension, (C) Pronunciation, (D) Scientific Vocabulary, and (E) Nonverbal Reasoning.

The aural comprehension section of the test is represented phonographically. The student listens to sixteen short passages and answers fifty multiple-choice items. The passages range in length from approximately thirty-three to two hundred and forty-one words. Norms based on several hundred subjects are available. The main criticisms of this test are, briefly: (1) Memory plays an important role in answering the questions correctly. (2) Familiarity with the subject matter may reflect in the scores, for example, "The movement of people from farms to the cities."

The second set of tests by the Educational Testing Service, The Inter-American Tests of Language, was published.

---

after its use in connection with a report on education in Puerto Rico. This test is composed of (1) Reading, (2) Language Usage, (3) Vocabulary and Interpretation in the Natural Sciences, and (4) Vocabulary and Interpretation in the Social Studies.

The Investigation of the Teaching of a Second Language

was conducted at the University of Chicago under a grant from the Rockefeller Foundation. A number of tests were made available, including (1) Elementary Structure, (2) Reading Comprehension, (3) Aural Comprehension, and (4) Writing Structure.

The examiner's part of the aural comprehension test is played from records and the student listens to the records and answers forty-four multiple-choice questions. There are four passages in the test lasting from four to five minutes each. The last passage is a dialogue recorded by two voices. The students may take notes and answer from nine to thirteen multiple-choice questions. The reliability of the aural comprehension test is given as, r = .81, split-half corrected by the Spearman-Brown prophecy formula. Norms are available based on the scores of 112 students.

A criticism of this test was made by one worker who reported that, without seeing or hearing the passages, he answered the questions which presumably were testing comprehension.

In multiple-choice items containing three choices each one would expect to guess about a third of them correctly. The scores for the four passages, however, were actually 80 per cent, 54 per cent, 78 per cent, and 58 per cent. These high scores confirmed my strong suspicion that familiarity with the areas of the passages would weigh heavily on the answers, thus invalidating the test as a measure of aural comprehension.


A team of two persons is necessary to give the U. S. Department of State's English Proficiency Examination. One person must be a native speaker of English, the other a native speaker of the other language. The test is designed to test the student's competence in English before he leaves his country. The first part of the test consists of eighty true-false questions which are read aloud to the student. In the second part of the test two anecdotes are read aloud to the student, and he writes answers to ten questions asked orally after each anecdote. In the third part the student writes a composition on a subject familiar to him. In the
last part he is interviewed to determine his ability to speak English. The true-false section of this test provides the only means for uniform scoring. No norms are available for this test.

"A Test of Aural Comprehension of English for Native Speakers of Spanish" is the title of a dissertation by Villarreal. This test was designed to measure the ability of native speakers of Spanish to understand connected passages of spoken English.

The experimental revision of this test consists of thirty-five connected passages of spoken English, twenty-five narratives and ten two-person conversations. The test materials were played to the subjects either by phonograph recordings or by magnetic wire-recordings. The subjects answered six and sometimes seven multiple-choice questions for each passage. Answer forms in Spanish were provided for Spanish-speaking subjects and answer forms in English for English-speaking subjects.

The experimental revision of the test was administered to 161 native speakers of Spanish and 104 native speakers of English. Concomitantly with the administration of the experimental revision, two ratings of the ability of each
Spanish-speaking subject to understand English were secured. The first rating was made by the subject himself and the second by a qualified observer; for example, the subject's teacher of English. Subsequently, the difficulty of each test passage was estimated on the basis of the percentage of successful responses yielded by both native and foreign subjects.

The reliability of Form A and Form B of the final test, as measured by the alternate form method, was estimated as, $r = .80 \pm .029$.\(^7\)

\(^7\) For 150 degrees of freedom an $r$ as small as .208 would be significant at the 1 per cent level of confidence.

There was a positive correlation between test scores and personal ratings of $r = .31$; between personal ratings and ratings by qualified observers, $r = .48$; and between test scores and ratings by qualified observers, $r = .78$. According to Villarreal, this result seems to permit the conclusion that self-ratings by foreign students on their ability to understand spoken English are undependable, but that comparisons of ratings by observers and test scores indicate strong evidence for the validity of this test.

Lado criticizes this test by saying:

The usual objection to the use of large passages for aural comprehension is in order with this test. The average length of the passages is 249 words; the first passage has 488 words; memory more than
aural comprehension must often be the deciding factor in giving the right answer.

8 Robert Lado, *op. cit.*, p. 64.

Tests developed at the English Language Institute, University of Michigan, include the *Examination in Structure*, the *Test of Aural Comprehension*, and the *English Language Test for Foreign Students*.


10 Robert Lado, "English Language Test for Foreign Students," *op. cit.*

There are three equivalent forms of 150 items each for the *Examination in Structure*. The student is allowed sixty minutes to finish this paper-and-pencil test and it can be scored objectively by an untrained person with the aid of a scoring key. Reliability of the test obtained by the split-half method corrected for length by the Spearman-Brown prophecy formula is, $r = .97$. Norms based on 357 students toward the end of training at the English Language Institute are available.

The *Test of Aural Comprehension* was constructed to measure achievement in connection with the teaching of English as a foreign language. The test has three equivalent forms: A, B, and C. The student indicates his understanding
of short passages of English by checking multiple-choice items on an answer sheet. The choices for the first part of the test are three pictures related to a short passage. The student chooses the picture that corresponds to the passage he hears and marks the appropriate place on his answer sheet. In the second part of the test the student chooses his answers from three phrases printed in English. The passages which the students hear range in length from three to forty-seven words. According to Lado, "The validity of the test rests in part on the fact that the items check specific matters of English vocabulary, pronunciation and structure in actual utterances."^11


Another aspect of the validity of this test was established by correlating the scores of 300 students enrolled in the English Language Institute with the composite opinions of eight teachers of the Institute who had been in daily contact with the students in class and at meals for a period of six weeks. Each teacher was asked to rate the students on a scale from zero to ten. The Pearson product-moment correlation was .85.\(^12\) The test-retest coefficient

\(^{12}\) Ibid.
half reliability for Form A corrected by the Spearman-Brown prophecy formula was, \( r = .88 \).

The test was also given to ten English-speaking students in a class on the teaching of English as a foreign language. Their average score was 97 per cent. Two students made perfect scores and the lowest score was 93 per cent.

One of the criticisms of this test centers around the undue importance of one or two key words in each sentence. In some cases the correct answer is decided by the sound of the plural ending of a word or the tense of the verb. In one case the correct answer is determined by the correct perception of the word "sinking." The stimulus sentence is "She is sinking." The three pictures to choose from represent a person sinking, a person singing, and a person thinking. (See Appendix A, Answer Booklet, Question 19). Although the undue importance of one or two words is pointed out as a weakness of the test, the ability to differentiate between the acoustical patterns of these key words undeniably is an important part of understanding a language.

The English Language Test for Foreign Students is a paper-and-pencil test. Scoring can be done by an untrained person with the aid of a punched scoring key. Part I of the test, "Structure," consists of forty multiple-choice
items. Each item consists of a sentence and three numbered phrases. The student is required to choose the phrases that are correct according to the sentence.

Part II, "Pronunciation," is composed of another forty items similar to the following example from the instructions for the test:

In each item you have three incomplete words, for example:

Example F: (1) Please open the -oor and come in.  
(2) Please si- down on that chair.  
(3) Is the win-ow open?

The sounds represented by - are the same in (1) door and (3) window. It is not the same in (2) sit. Mark "1.3" on the answer sheet, like this:

\[
\begin{array}{cccccc}
F. & (\_\_) & (\_\_\_) & (\_\_) & (\_\_) & (\_\_\_) \\
1.2 & 1.3 & 123 & 2.3 & 0 \\
\end{array}
\]

An additional fourteen items under Part II are similar to the following example from the instructions for the test:

Example K:

PEN-CIL

You may write with a pencil.

1 2 1-PEN receives the heaviest stress. Mark "1" on the answer sheet, like this:

\[
\begin{array}{cccccc}
K. & (\_\_\_) & (\_\_) & (\_\_) & (\_\_) & (\_\_) \\
1.2 & 1.3 & 123 & 2.3 & 0 \\
\end{array}
\]

The results obtained from this test were correlated with the Test of Aural Comprehension, the Test of Aural Perception, and the Examination in Structure. The correlation
was given as, $r = .89$. A more complete explanation of how this correlation was determined was not given. The reliability was reported as, $r = .95$ for forty-one cases obtained by the split-half method corrected by the Spearman-Brown prophecy formula.\(^\text{13}\)

\(^{13}\) Robert Lado, *English Language Test for Foreign Students*, op. cit.

2. STUDIES WHICH TREAT THE RELATIONSHIP BETWEEN SPEECH-SOUND DISCRIMINATION AND ARTICULATORY PROFICIENCY.

Lundeberg and Tharp\(^\text{14}\) were perhaps the first to attempt to test the ability to recognize and discriminate the significant sounds of language in a test of second language learning.

A *Test of Aural Perception*\(^\text{15}\) constructed by Lado consists of examiner's materials, an answer sheet, and a scoring sheet or stencil. It is made up of 100 items in which the student is asked to listen to three or more words or sentences and to indicate which, if any, are the same. According to Lado, "The validity of this test is based on the linguistic analysis of what is tested and how it is
tested. Each one of the problems is operative in the language as proved by phonemic analysis . . ."16

16 Ibid., p. 63.

Thirty-seven scores on the test of Aural Perception were correlated with scores made by the same students on the test of Aural Comprehension, \( r = -0.14 + 0.16 \). Aural Perception scores and total scores on the Seashore Musical Aptitude test, Form A, were computed for twenty-nine students. The correlation was, \( r = 0.28 + 0.17 \) and was not significant at the 5 per cent level of confidence.17

17 A correlation of .16 would not be significant at the 5 per cent level of confidence.

Reliability was reported as, \( r = 0.90 + 0.02 \), test re-test.

A number of studies have been conducted relative to the sound discrimination ability of normal speakers of English and speakers of English with functional speech disorders. Although the results of the studies were conflicting, the investigations were conducted because of the assumption that sound discrimination is an important part in the ability to speak properly.
In 1931, Travis and Rasmus investigated the sound discrimination ability of 383 normal American speakers and 165 American speakers having functional articulatory defects. The subjects, who ranged from kindergarteners to college freshmen, were asked to tell whether pairs of spoken nonsense syllables were the same or different. The results of this study showed that the subjects having articulatory defects missed a significantly greater portion of syllabic discriminations than did those persons who had no articulatory errors in their speech. In this study there were 366 pairs of sounds of which 331 pairs were different and thirty-five pairs were alike. Although consonant-vowel pairs were included in the test, consonants and vowels were not compared in a syllable pair. The results of this study did not demonstrate a significant correlation between scores on the Travis and Rasmus speech-sound discrimination test and intelligence as measured by the Stanford-Binet test.

Carrell studied the failure of people with normal hearing but with articulatory defects and people with normal speech and normal hearing to discriminate between pairs of sounds. His methods were similar to those of Travis and Rasmus but differed from their methods in the controls he
imposed. His paired sounds were recorded on phonograph discs by the same speaker throughout, and an attempt was made to keep intensity, pitch, and inflection constant. The subjects were of elementary school age. Although Carrell's results were not as conclusive as those of Travis and Rasmus, he found a positive relationship between speech-sound discrimination and articulatory proficiency.

In 1938, Hall\textsuperscript{21} attempted to determine the auditory factors in functional articulatory defects. In one part of her study she investigated the relationship between speech-sound discrimination ability and articulatory defects of elementary school children. In the second part of her study she used as subjects college freshmen ranging from marked articulatory defectives to superior speakers. In addition to having her subjects discriminate between paired syllables, she read nonsense syllables to them and then asked them to identify the first pattern among five nonsense phrases. The first pattern was read as one of five nonsense phrases. In this study, no significant difference was found between the normal speakers and speakers with defective articulation in the ability to discriminate between pairs of syllables and nonsense syllable phrases.

\textsuperscript{21} Hall, \textit{op. cit.}
Hansen investigated the application of sound discrimination tests to functional articulatory defectives with normal hearing. Hansen stated that although speech-sound discrimination tests involving gross sound differences may be suitable for children, a test requiring finer degrees of differentiation is necessary to detect the differences among abilities of adults. He used three tests of sound discrimination: The Seashore Measure of Timbre, the Travis-Glaspy Sound Discrimination Test (sic), and a test of vowel sound discrimination. The experimental subjects were students with functional articulatory defects who had had previous clinical therapy, students with functional articulatory defects who had not received clinical therapy, and normal speakers. He found no significant differences among the groups.

Templin investigated the speech-sound discrimination ability of elementary school children. Her first objective was to determine whether there was any variation in the ability of children to discriminate between identical and unlike syllables when the position of the discriminative
element was changed from the initial to the medial or final position of the syllable. The second objective was to determine whether a short test of sound discrimination could be used satisfactorily.

In order to determine whether the position of the discriminative element was important, two tests, similar to the Travis-Rasmus test of Sound Discrimination, were constructed. The two tests, each containing 100 items, were essentially the same except for the position of the consonant element. In the first test the consonant was in the initial position. In the second test the consonant was in the second position. These two tests were administered to 301 elementary school children in grades two through six. The results of her study indicated a significant difference in scores when the discriminative element changed from the initial to the medial or final position. The test items in both tests that had been missed by approximately 10 per cent of the total group, or passed by only 10 per cent, were discarded. The seventy-middle items were used as a shorter test. The split-half reliability coefficient corrected for length by the Spearman-Brown prophecy formula for the short test was .83.

Reid,25 in studying factors relating to functional

25 Gladys Reid, op. cit.
articulatory defects, used the Templin Short Test of Sound Discrimination. She found the highest correlation among factors tested, between the speech-sound discrimination scores and the speech ratings. The Pearson product-moment correlation was $0.37 \pm 0.09$. This was significant at the 5 per cent level of confidence. Her subjects were thirty-eight elementary school children.

Anderson\textsuperscript{26} studied thirty-one children from kindergarten through the fourth grade who had been diagnosed as having defective articulation of the consonant $\{s\}$ to determine the relationship of normal and defective articulation of the consonant $\{s\}$ in various phonetic contexts to the auditory discrimination between normal and defective production. She administered a picture articulation test containing the consonant $\{s\}$ in various phonetic contexts and a speech sound discrimination test designed to test discrimination of good and poor $\{s\}$ productions. She found a correlation of $0.25$ between the number of misarticulations of the consonant $\{s\}$ and the total number of discrimination errors, and a correlation of $0.38$ when the discrimination in the particular combination in which $\{s\}$ was misarticulated was compared with the frequency of the $\{s\}$ errors. The correlations were not statistically significant.

\textsuperscript{26} Patricia Ward Anderson, \textit{op. cit.}
3. STUDIES WHICH DEAL WITH MEASURES OF INTELLIGIBILITY

In the literature on foreign language teaching, methods are described for the measurement of what is variously called "pronunciation," "oral production," and "intelligibility."

In an Investigation of Second Language Teaching,\textsuperscript{27}

\textsuperscript{27} Agard and Dunkel, \textit{op. cit.}

tests of "Oral Production" are described. These tests were designed to measure three "Essential aspects of oral skill: (1) the ability to report a single, simple act or situation in precise words, (2) the ability to express a sequence of ideas fluently, (3) the ability to converse. . .The major criterion in these three aspects was intelligibility.\textsuperscript{28}

\textsuperscript{28} Ibid., p. 56.

In the directions for part II, designated as "sustained speech," the examiner is requested to rate each of four qualities on a three point scale: Fluency, Vocabulary, Pronunciation and Enunciation, and Grammatical Correctness. The description of the scale for Pronunciation and Enunciation was as follows:

\begin{itemize}
    \item \textbf{2} - Sufficiently approaches native speech to be completely understandable.
    \item \textbf{1} - Can be understood, though with difficulty, because there are sounds which he does not utter correctly.
\end{itemize}
0 - Would not be understood by natives because his pronunciation is so different from theirs.\(^{29}\)

\(^{29}\) Ibid., p. 58

Agard and Dunkel state that judgments for Part II were very subjective, "... useful models for the production of sustained speech could not be efficiently supplied; thus, despite the rating scale supplied, examiners necessarily relied on the standards they were accustomed to set for their own students with respect to fluency, vocabulary, pronunciation, and grammar."\(^{30}\)

\(^{30}\) Ibid., pp. 99-100.

Lado describes a direct method of testing pronunciation whereby the student is shown a series of pictures and is asked to tell what he sees.

I used pictures as a stimulus as follows: The student is presented with a series of pictures and is asked to tell what he sees. He thinks he is being tested on vocabulary and grammar because the pictures are arranged as "actions," "things," "qualities," "series," etc., but the examiner listens for pronunciation instead. To test the mastery of /d/ between vowels for example there is an "action" picture of a man reading and a "quality" picture of a big ladder and a little one together. The examiner listens only to the top quality of the /d/ in reading and ladder. If he is not trained in phonetics he listens only for reading versus wheathing and for ladder versus lather. In about twelve minutes he can sample 35 pronunciation problems twice.\(^{31}\)

This method was called a **Test of Oral Production**. Lado\(^{32}\)


reports that the product moment correlation between the Test of Aural Perception and the Test of Oral Production was \(0.30 \pm 0.14\).\(^{33}\)

\(^{33}\) This correlation would be significant at the 1 per cent level of confidence if based on the scores of forty or more subjects.

In the past twenty-five years a considerable amount of information has been gathered relative to the measurement of intelligibility. These studies, however, have not been directly concerned with the measurement of the speech proficiency of foreigners.

This approach to the measurement of speech proficiency is defined thus: "Intelligibility: Per cent intelligibility of a communication system is the percentage of the speech units spoken by a talker or talkers that is understood correctly by a listener or listeners."\(^{34}\)

Wolf was possibly the first to publish an account

of the differences in the intelligibility of sounds and words. He and his colleagues spoke various sounds and noted the distances at which it was possible to distinguish the sounds. He noted that vowels could be understood at greater distances than consonants and that sounds could be heard at greater distances in a concert-hall when it was empty than when it was full of people.

In 1929, Fletcher and Steinberg published a report

of quantitative methods for the measurement of speech transmission. Lists of syllables, words, and sentences were formulated. Readers spoke the sentences and syllables over various transmission systems and listeners wrote down what they thought they heard. Articulation scores were assigned on the basis of the proportion of the listener's correct responses.

The methods developed at the Bell Telephone Laboratories were designed primarily to measure the performance of transmission systems. Jean Brady Jones, possibly, was the

First to have a principal objective in the measurement of the intelligibility of the speaker. Her "Enunciation Test" provided for testing students in groups of eleven. Ten of the listeners sat at right angles to the speaker while one of the students read a twenty-word list. The listeners wrote down what they heard. Scores were based on the proportion of right responses.

During World War II, interest in speech intelligibility was greatly increased. At the Psycho-Acoustic Laboratory, Harvard University, lists of fifty English monosyllabic words were developed. These lists were called phonetically balanced because of the attempt that was made to have the consonants occur in a list in approximately the same proportion that they occur in English. By altering the order of the words the lists could be used indefinitely with the same listeners. As before, the words were read by a speaker and the responses written down by a panel of listeners.

About the same time the work at Harvard was being carried on, a voice-communication training program was being...
developed at the Voice Communication Laboratory, Waco Army Air Field, Texas. In order to assess improvement in voice communication, Haagen developed twenty-four word write-down tests. Haagen selected the words from the most frequently used words in Thorndike's list. The lists were equated on the basis of the reception scores for the individual items when spoken and heard over an airplane intercommunication system in the presence of simulated cockpit noise.

In order to overcome the disadvantages and labor inherent in scoring write-down intelligibility tests, Haagen constructed a second test of twenty-four equated lists, each composed of twenty-four multiple-choice items. The wrong choices that were grouped with an item on the multiple-choice answer sheet were the three words that had been most frequently written in error by panels of listeners who had heard the test word called by at least twelve speakers.
The multiple-choice tests were found to be satisfactory in these respects: twenty-four lists of twenty-four words each were not significantly dissimilar in mean difficulty or variability as tested by analysis of variance; the test satisfactorily distinguished between trained and untrained groups of personnel; the tests were internally consistent to the extent that $r$, split-half, corrected for length was .77 to .85 ($N$, approximately 200); the consistency of items was indicated by $r = .89$ from the administration of the test from one panel to a second.

The tests were also subjected to an item analysis at the Voice Science Laboratory, Purdue University, and again were found to be equated in mean difficulty and variance.

4. STUDIES RELATED TO INDIRECT METHODS OF ASSESSING ARTICULATION PROFICIENCY.

The assumption basic to an indirect method of assessing articulatory proficiency is that what a person thinks he says or his stated choice of controlled speech samples is related to the way he talks or to his general competence in the language.
Indirect methods of testing articulation are not new. Lundeberg and Tharp were among the first to test the proficiency of students studying French and Spanish as a foreign language. The English Language Test for Foreign Students includes a section on pronunciation. The student is presented with items which consist of pairs of words. The student must decide whether or not the words rhyme. In the remaining items the students check the syllable that receives the loudest stress in words. Examples of items for this part of the test were presented previously in this chapter under "Studies which deal with the testing of comprehension of English." In describing this method, Lado writes,

To check the relation between what the student records as his answer and what he actually says in reading the test aloud I gave a forty-item test of this kind to 17 foreign students and then without having looked at their answers I asked each one to read the test aloud while I scored their pronunciation on it. The correlation on the 40 item test was .85. If we increase the length of the test to 80 items the correlation would be expected to rise to .92. Such a correlation justifies the use of the technique where the more direct ones cannot be used satisfactorily.

Lundeberg, op. cit.

Lado, op. cit.

Griffith investigated the relationship between phonetic patterns of American college students with inferior, average, and superior articulation and controlled speech sound stimuli. Ruling out defective hearing, the question was asked whether errors in sound discrimination were attributable to a preference on the part of the individual to sound patterns similar to his own.

To answer this question, 253 subjects recorded nonsense syllables containing four consonant sounds and were judged as to their articulatory excellence. From this group, subjects who were rated as superior, inferior, and average in articulation ability listened to nonsense syllables similar to those mentioned above that had been previously evaluated in terms of excellence. Statistical analysis failed to reject the hypothesis: "Errors in sound discrimination are attributable to a preference on the part of the individual to sound patterns similar to his own."

5. STUDIES WHICH TREAT PREDICTION OF LETTERS.

Shannon selected a short passage unfamiliar to the
person who was to do the predicting. The subject was asked to guess the first letter in the passage. If the guess was correct, he was so informed, and proceeded to guess the second letter. If the guess was not correct, he was told the correct letter and proceeded to guess the second letter. In one such experiment Shannon found that for a total of 129 letters, eighty-nine (or 69 per cent) of the letters were guessed correctly on the first attempt.48

48 Ibid., p. 6.

In an extension of the above experiment the subject was required to guess the letters as above, but he was told if he had guessed wrong and was asked to continue to guess until he found the correct letter. He was informed when he found the correct letter and proceeded to guess the second letter in the same manner. In one such experiment, out of 102 symbols, the subject guessed seventy-nine letters correctly on the first guess, eight letters on the second guess, three on the third guess, two on the fourth and fifth guesses. Only eight letters required more than five guesses.49

49 Ibid., p. 7.

In order to determine how predictability depends on the number $N$ of preceding letters known to a subject, a more
involved experiment was carried out. One hundred samples of English text were selected at random from a book, each fifteen letters in length. The subject was required to guess the text, letter by letter, for each sample as in the preceding experiment. "To aid in prediction the subject made such use as he wished of various statistical tables, letters, digram and trigram tables, a table of the frequencies of initial letters in words, a list of the frequencies of common words and a dictionary."50

50 Ibid., pp. 7-8.

The data obtained from the subject on the basis of the 100 samples were summarized in a table. The entries in the columns of this table corresponded to the number of preceding letters known to the subjects plus one; the row was the number of the guess. The entry in column N at row S was the number of times the subject guessed the right letter at the Sth guess when (N-1) letters were known. He found that the prediction gradually improved with increasing knowledge of the past (preceding letters known to the subject) as indicated by a large number of correct first guesses and the smaller numbers of high rank guesses.51

51 Ibid., p. 9.
Summary of the Chapter. This chapter summarized the studies in the literature relevant to this investigation. The studies were described under five categories:

1. Studies that treat the testing of comprehension of English as a foreign language.
2. Studies that treat the relationship between speech-sound discrimination and articulatory proficiency.
3. Studies that treat measures of intelligibility.
4. Studies that treat indirect methods of assessing articulatory proficiency.
5. Studies that treat the prediction of the letters of English.

The next chapter is concerned with the selection and preparation of materials employed in this study, the selection of the experimental subjects, and the procedures followed in testing the subjects.
CHAPTER III

METHODS AND PROCEDURES FOR SELECTION AND PREPARATION OF THE MATERIALS, SELECTION OF SUBJECTS, AND TESTING OF SUBJECTS

This study was designed to investigate four null hypotheses related to the learning of English as a foreign language. These hypotheses were:

(1) There is no difference between native and non-native speakers of English (a) in the ability to predict the letters of English, (b) in the ability to comprehend English, (c) in the ability to differentiate between the sounds of English, (d) in intelligibility when speaking English, and (e) in the ability to judge controlled patterns of speech.

(2) There is no difference between native speakers of Spanish and Japanese in the variables that are listed in 1 immediately above.

(3) There is no relationship among the variables that are listed in 1 immediately above.

(4) Judgments of controlled speech patterns are not reliable or valid estimates of intelligibility.

Six measures were selected to test these hypotheses. Because instruments for testing were unavailable or unsuitable for certain aspects of this investigation, three of the tests were designed especially for this study. Part I of this chapter will pertain to the criteria for the selection
and preparation of each of the six tests and a description of each of the tests that are selected or an explanation of the steps followed in the preparation of the materials. Part II will state the criteria for the selection of the experimental subjects and a description of the subjects who were chosen for this study. Part III will pertain to the controls relative to the testing procedures.


Comprehension of English: Aural Comprehension.
The Test of Aural Comprehension, Form A, was selected to measure aural comprehension. The following criteria supported this selection: (1) The test was designed to assess comprehension of spoken English. (2) The test was designed for use with non-native speakers of English. (3) The test had been validated and found to be reliable. (4) The test is relatively easy to administer and score. (5) The test requires not more than one-half hour for administration. (6) The test was designed for group testing.

The Test of Aural Comprehension was constructed at the English Language Institute, University of Michigan, to determine a foreign student's competence in English and to measure achievement in connection with the teaching of English as a foreign language. The test has three equivalent
forms—A, B, and C. The student indicates his understanding of short passages of English by checking multiple-choice items on an answer sheet. The choices for the first part of the test are three pictures related to a short passage. The student chooses the picture that corresponds to the passage he hears and marks the appropriate place on his answer sheet. In the second part of the test the student chooses his answers from three phrases printed in English. The passages which the students hear range in length from three to forty-seven words.  

2 Appendix A contains a copy of the examination booklet and an answer sheet for this test.

According to Lado, "The validity of the test rests in part on the fact that the items check specific matters of English vocabulary, pronunciation, and structure in actual utterances." In addition, the scores of 300 students enrolled in the English Language Institute were correlated with the composite opinions of eight teachers of the institute who had been in daily contact with the students in class and at meals for a period of six weeks. Each teacher was asked to rate the students on a scale from zero to ten. The Pearson product-moment correlation was .85. The test-retest coefficient of reliability (Forms A and C) was .87.
The split-half reliability for Form A corrected by the Spearman-Brown prophecy formula was .88.\(^4\)

\(^4\) Ibid.

The test, including instructions, can be given to more than one person at a time and required approximately twenty-five minutes for administration. The directions accompanying the test of aural comprehension provide instructions for the students and are printed on the front of the examination booklet. (See Appendix A). The direction for the examiner are as follows:

The examiner will read the following material to the students, allowing enough time after each item for them to record their answers. The length of the pause should be between 12 and 15 seconds. The examiner should read in a clear natural style; no effort should be made to increase or decrease the difficulty of each item.\(^5\)

\(^5\) See Appendix A "Examiners material" for the complete instructions and stimulus sentences.

The instructions and sentences were recorded on a Bell Re-Cord-O-Fone Model RT-65 B tape recorder-reproducer at seven and one-half inches per second tape speed. The reader, an adult female native speaker of English, was seated in a sound-treated room. She spoke into a Shure Brothers Model 55 Unidyne microphone. The speaker maintained a constant distance from the microphone with the aid of a wooden stick attached to the microphone. A stop watch was
used by the experimenter to determine the time between stimulus passages. A light was flashed to signal the reader at the beginning of each passage.

Speech-sound Discrimination. As mentioned in Chapter I, part of this investigation was concerned with the relationship between the intelligibility of the non-native speaker of English and his ability to discriminate between the sounds of English. Templin's "Short Test of Sound Discrimination,"

6 Templin, op. cit.

was selected to determine the sound-discrimination ability of the subjects who took part in this study. The following criteria supported this selection: (1) The test is reasonably valid and reliable. (2) The test is relatively easy to administer and score. (3) It requires not more than ten minutes to administer.

Templin's "Short Test" is a modification of the Travis-Rasmus Test of Sound Discrimination. 7 Whereas the

7 Travis and Rasmus, op. cit.

Travis-Rasmus test contained 366 pairs of sounds, Templin's test contains seventy pairs. Templin tested the validity of her test by obtaining the educational age, reading comprehension age, and spelling age from standardized tests, in addition to grade level and hearing acuity for 243
children in the third, fourth, fifth, and sixth grades. The number of errors on the Short Test was correlated with these measures. The obtained correlations were:

<table>
<thead>
<tr>
<th>Measure</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational Age</td>
<td>-.207 ± .04</td>
</tr>
<tr>
<td>Reading Comprehension Age</td>
<td>-.189 ± .04</td>
</tr>
<tr>
<td>Spelling Age</td>
<td>-.260 ± .04</td>
</tr>
<tr>
<td>Grade</td>
<td>-.129 ± .05</td>
</tr>
<tr>
<td>Hearing Loss</td>
<td>-.013 ± .06</td>
</tr>
</tbody>
</table>

A correlation of .138 is significant at the 5 per cent level of confidence for 200 degrees of freedom. Thus, three of the correlations above are significant at the 5 per cent level of confidence.

The split-half reliability coefficient of correlation was reported as, r = .72. Corrected for length by the Spearman-Brown prophecy formula, the correlation was, r = .83.

Administration time for this test amounted to approximately ten minutes, including instructions. Scoring is relatively simple in that the test can be scored with the aid of a key. A score for an individual is determined by subtracting the number of errors from seventy (the total number of pairs of syllables). In order to keep the stimulus material the same from situation to situation, the syllable-pairs were recorded on a Magnecorder type PT6-PS recorder at seven and one-half inches per second tape speed. While recording, the examiner held a Western Electric Model 633 microphone at the corner of his mouth. Sound pressure level was monitored visually on the W meter on the face of the Magnecorder. The directions for the subjects were also
recorded by the experimenter. A copy of the instructions was given to each subject to follow while the directions were being played during the testing situation. Each subject was provided with an answer sheet containing seventy-five numbered spaces. Although the subjects heard only seventy pairs of syllables, they were not told how many pairs they would hear. The directions read to the subjects were as follows:

You will hear pairs of syllables. If you think the syllables in the pair are the same, write S on your answer sheet. If you think the syllables in the pair are different, write D on your answer sheet. For example, you may hear (te--de). If you think both of the syllables are the same you will write S on your paper. If you think they are different, you will write D on your paper.

Intelligibility. Form B of the twenty-four word multiple-choice test developed at the Voice Communication Laboratory, Waco, Texas, was selected to measure intelligibility in this study. Form B consists of twelve equated lists of words, each comprised of twenty-four words. In the usual procedure for the administration of the test the
speaker reads one of the lists of words to a panel of at least ten listeners. The listeners are provided with a multiple-choice answer sheet and draw lines through the words they think they hear. The wrong choices that were grouped with an item on the multiple-choice answer forms were the three words that had been most frequently written in error by panels of listeners who had heard the words spoken by at least twelve speakers. A constant distance is maintained between the speaker and listening panel and a constant-level noise barrier is introduced into the speaker-listener system. As soon as one person has spoken he takes his place as a listener and one of the listeners becomes a speaker.\footnote{C. H. Haagen, "Intelligibility Measurement: Techniques and Procedures used by the Voice Communication Laboratory," \textit{op. cit.}}

Only one third of the subjects in this study were native speakers of English. If non-native speakers of English acted as listeners, as above, the reading ability of the non-native judges might influence the scores of the speakers. For the purpose of this study, therefore, the responses of all subjects were recorded as the lists of words were read for the intelligibility test. These responses were subsequently played to panels of native speakers of English.
The subjects spoke into a Western Electric 633 microphone. Their responses were recorded on a Bell Record-O-Fone Recorder-Reproducer Model RT-65B Tape recorder at seven and one-half inches per second tape speed. No attempt was made to control the sound-pressure level of the speaker's responses. The recording level of the tape recorder was not changed during the recording sessions.

The tapes containing the responses of the sixty speakers for the intelligibility test were played to five groups of students who were enrolled in Speech 401 (Principles and Practice of Effective Speaking I) during the Autumn Quarter, 1952. Each group of twenty listeners judged the responses of twelve speakers. The judges were provided with the answer sheets for Form B of the Waco, twenty-four word multiple-choice intelligibility test. The tapes were played at a constant level on a Bell Record-O-Fone Model RT-65B Recorder-Reproducer connected to an Altec Speaker Model 400B mounted in a baffle.

Noise was introduced into the listening environment by means of a recording of white noise played on a Presto recorder-reproducer Type K-8-M. The noise measured 60 db on a General Radio Company Sound Level Meter Type 759-B (C-scale) fifteen feet from the speaker of the Presto recorder-reproducer. The Presto speaker and the Altec speaker were placed side by side in front of the room fifteen feet from
the nearest listeners. The listeners sat in rows of chairs parallel to the front of the room.

Comprehension of English: Letter Prediction.

Methods employed by Shannon\textsuperscript{12} to measure the ability to predict the letters of English were described in Chapters I and II. His subject was tested individually and was informed verbally when the correct letters were guessed. A special test was designed for the purposes of the present study to permit group testing of the students without the necessity of someone telling them when the correct letters were found.

Newspaper text was selected as the source of the material to be predicted by the subjects, as it was assumed that newspaper text would be the most common source of English relatively familiar to all subjects taking part in this study. Researchers at the Naval Air Station, Pensacola, Florida, in preparation for a study of the predictability of two different vocabularies, newspaper and flight patter, had selected sixty-nine five-syllable phrases, textually connected, from newspapers. The words were taken from the first five syllables of the fifth line of successive news stories. If a proper noun was involved, the immediately successive material was used. A further restriction was

\textsuperscript{12} Shannon, \textit{loc. cit.}
observed in that each phrase had to contain three or more 
words. Contractions were avoided.

From the above sixty-nine phrases six passages con­
taining at least fifteen letters each were chosen for the 
present experimentation. The passages were:

1. in and around the 4. who have been done
2. on the east fringe 5. was taken because
3. here at a meeting 6. to decide whether

Only the first fifteen letters of each phrase were used in 
the test.

Answer sheets which contained fifteen columns of 
twenty-seven symbols each were prepared (See Appendix D). 
Each column was composed of the English orthographic alphabet, 
A through Z, plus (-). The latter symbol was included to 
indicate the space between words.

The answer sheets were coded by daubing the letters 
of the message with a solution of potassium thio-cyanate 
and water. When the solution dried, no clue to the identity 
of the treated letters could be observed. For example, for 
the message "in and around t" the letter i in the first 
column was daubed with the chemical solution, the letter n 
in the second column, and the (-) in the third column. This 
procedure was followed for the remaining letters of the 
message. In the testing situation a bottle containing a 
solution of green ink and ferric sulfate was provided each
subject. When the green ink was applied to a letter on the answer sheet, the ink remained green unless it was previously coded with the potassium thio-cyanate solution. In that case the ink immediately turned brown. The instructions for the subjects were:

You are to guess the messages on the yellow answer sheets in the following way:

Turn to the sheet marked "P". Perhaps you guess that the first letter of the message is Y. Make a mark on the letter Y in the first column with the pen and ink that has been provided. If it is correct, the green ink will turn brown. If you try it, you will see that it does not turn brown. Perhaps your second guess is that the letter is E. Make a mark on the letter E. The ink remains green so you have not found the correct letter. Keep trying until you find the right answer. In this case the correct answer is 3. If you make a mark on 3 you will see that the ink turns brown. Write the letter in the blank provided above the column. Count the number of guesses and write that number above the letter. Then proceed to the next column to find the second letter. Begin by marking over the letters in the second column and continue until you find the right letter. By writing in the row marked (-) you will be able to find out if the letter or letters you found compose a word. Thus, (-) indicates a space between words.

Try to get as many letters as possible on the first guess. There is no time limit for this test. Take as much time as you wish. You should, however, be able to complete each page in about ten minutes.

The last word of the message may or may not be complete. The first word is always complete, that is, you will be able to find all the letters for the first word. Please begin.

The answer sheets were marked "1" through "6." Every answer sheet "number 1" was coded by means of the chemical solution with the message "in and around t;" answer sheets
"number 2" were coded with the message "on the east fri,"
and so on for the six phrases. In order to avoid an order
effect, the answer sheets were systematically rearranged
following a Latin square technique. Thus, the first set
contained phrases 1, 2, 3, 4, 5, 6; the second set 2, 3, 4,
5, 6, 1; the third set 3, 4, 5, 6, 1, 2; and so on. A page
of instructions and a practice answer sheet were attached
to the front of each set of answer sheets. The practice
answer sheet (P) was coded with the message "so you w."

Controlled Speech Patterns. Direct testing of speech
usually requires that a sample of a subject's speech is
judged either by a single judge or a group of judges. The
important consideration in direct testing of pronunciation
is that the subject says something. The assumption under­
lying indirect testing of pronunciation is that an index of
an individual's proficiency in speech can be obtained
without requiring the subject to talk. An implication of
this assumption is that a person's standard of speech ex­
cellence is in part determined by his own speech. If his
speech is good, he will tend to be a competent judge of the
speech of others. If his speech is relatively poor, he will
tend to be a poor judge of the speech of others.

To test the assumption that a person's standard of
speech is related to the way he talks, stimulus material
was prepared for presentation to the experimental subjects.
The stimulus material consisted of recorded speech patterns of native and non-native speakers of English. One form of the stimulus material contained passages of speech that varied in pronunciation excellence from very poor speech to very good speech. The ranks for the passages were based on the judgments of 100 college students. It was assumed that listeners and especially the non-native speakers of English would deviate in ranking these passages from the ranks of the criterion judges, and that the obtained deviation scores would be related to intelligibility scores.

The second form of the stimulus material contained pairs of four-word phrases. One member of each pair had been previously judged to be better than the other. It was assumed that non-native speakers would not be able to indicate the "better" phrases as often as native speakers and that error scores would be related to scores obtained for intelligibility. The next section of this chapter is devoted to a description of the steps in the preparation of the stimulus material.

**Stimulus Material for Controlled Speech Patterns.**

In the preparation of the stimulus material for controlled speech patterns, two factors were considered. First, an international student may find that a given word is unfamiliar to him and not know how it should be pronounced. In order to reduce this possibility and to insure that the
student has heard the word many times even though he has had little experience with English, words were selected that occur frequently in spoken English. Second, a student might rate a given passage favorably because he understands the meaning of the passage or unfavorably because he does not understand the meaning of the passage. To reduce the importance of meaning per se, short (four-word) phrases were employed in the tests instead of sentences.

The stimulus material selected for controlled speech patterns was the five-syllable phrases equated by Walker and Black\(^{13}\) for rate and intensity. Only phrases containing four or more words were retained. All phrases containing five words were reduced to four by eliminating the first word of each phrase. The words of the remaining phrases were compared with the words in Sections 1 through 4 in Ausherman and Black's study in which the frequency of occurrence of words in the formal spoken vocabulary of college students was determined.\(^{14}\) Sections 1 through 4


include 1,297 of the more frequently used words of the 6,864 different words enumerated in the study. If any word had a higher rating than 1/2, the phrase in which that word appeared was eliminated. All remaining phrases were checked for duplications, leaving 138 four-word phrases. Of these phrases, 135 were used in the next portion of the study.

**Speakers.** A total of fifty-four adult speakers, twenty-nine males and twenty-five females each read ten of the phrases described above. Thirty-one of the speakers were non-native speakers of English. Sixteen of the speakers spoke English as a native language. Fourteen of the native English speakers were normal speakers without speech defects. In order to have a greater variety of examples of good and poor speech spoken by native speakers of English, two subjects were selected who had speech defects and seven of the normal speakers read a second set of phrases with a pencil held between their teeth.

The breakdown of the speakers in terms of their native country is shown below:

<table>
<thead>
<tr>
<th>Country</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>16</td>
</tr>
<tr>
<td>Japan</td>
<td>5</td>
</tr>
<tr>
<td>Thailand</td>
<td>4</td>
</tr>
<tr>
<td>Germany</td>
<td>3</td>
</tr>
<tr>
<td>Latvia</td>
<td>3</td>
</tr>
<tr>
<td>Hawaii</td>
<td>2</td>
</tr>
<tr>
<td>Turkey</td>
<td>2</td>
</tr>
<tr>
<td>Slovakia</td>
<td>1</td>
</tr>
<tr>
<td>France</td>
<td>1</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1</td>
</tr>
<tr>
<td>Pakistan</td>
<td>1</td>
</tr>
<tr>
<td>Italy</td>
<td>1</td>
</tr>
<tr>
<td>Iraq</td>
<td>1</td>
</tr>
<tr>
<td>Peru</td>
<td>1</td>
</tr>
<tr>
<td>Bolivia</td>
<td>1</td>
</tr>
<tr>
<td>China</td>
<td>1</td>
</tr>
<tr>
<td>Norway</td>
<td>1</td>
</tr>
</tbody>
</table>
Recording the Phrases. The speakers were seated in a sound treated room one at a time. A wooden stick seven inches in length was attached to a Shure Brothers Model 55 Unidyne Microphone. Each speaker was instructed to place the center of his forehead against the end of the stick, in order to assure that the distance between the lips and microphone would remain constant for all speakers. The microphone of a Decibeloscope was placed three feet away from the speaker. The speaker monitored his sound pressure level by watching the lights on the Decibeloscope. A minimum sound pressure level of 55 db caused a green lamp to light; a maximum sound pressure level of 70 db caused the red, or second, lamp to light.\footnote{The decibeloscope was calibrated by sustaining vowels at 55 db and 70 db according to measurements on a General Radio Sound Level Meter (C-scale) and setting the lights to light at these intensities.} The lights were set at the maximum speed settings. A thyrotron controlled reading light over the right shoulder of the speaker dimmed when the speaker's sound pressure level exceeded 70 db. The speaker was instructed to light the green light when speaking. He was to regard the lighting of the red light or the dimming of the reading lamp as a warning that he was speaking louder than necessary.

The speakers were requested to speak as distinctly as possible and were instructed to take a deep breath and
use the same breath to speak the ten phrases. A buzzer was sounded to signal the start of each phrase. The first speaker read phrase one to ten, the second speaker phrase five to fifteen, etc. Thus, every phrase was spoken by two different speakers. The recording was done on a Magnecorder tape recorder Type PT6-PS.

Construction of Tape A and Tape B. The sound of the buzzer was edited from the tape that contained the responses of the fifty-four speakers. The ten phrases for each speaker were abutted. The first interval of nine seconds was measured for each speaker and the remaining tape removed.  


Lewis and Sherman in their study concerned with the measurement of the severity of stuttering by means of scaled passages found that opinions of 40 judges on sample length indicated that six-second samples were too short, ten-second samples were about right, and fifteen-second samples were too long. Lewis and Sherman adopted nine seconds as the length of the sample to be used.

Each nine-second passage was connected by a strip of timing tape four seconds in length. Each nine-second passage contained at least five four-word phrases. The edited tape was reproduced on a Bell Record-O-Fone Model RT-65B tape recorder at seven and one-half inches per second tape speed. Hereafter, the latter tape will be referred to as "Tape A."
The first five phrases of each of the original nine-second passages were then separated by a four-second section of timing tape. The phrases were systematically rearranged so that no two phrases spoken by the same speaker directly followed each other. This resulted in a tape composed of 270 four-word phrases, each phrase separated by a four-second pause. This tape was then reproduced as above. Hereafter, this tape will be referred to as "Tape B."

**Assignment of Scale Values to the Tapes A and B.**

Tapes A and B described above were played on a Bell Record-O-Fone recorder reproducer Model RT-65B to 100 students enrolled in Speech 401 at The Ohio State University during the Summer Quarter, 1952. Of these students, thirty-six were females with a mean age of 21.7 years and sixty-four were males with a mean age of 24.5 years. All the students were native speakers of English. This group will be called Group A.

The students were given a copy of the following instructions for Tape A.

> You will hear a number of people speaking phrases. You are asked to rate these people in terms of how well they speak these phrases. Do they say the words the way you think they should be said in English?

> You are to rate them on a scale from 1 to 8.

> 1 means the person speaks without error. The words are spoken in the way they should be said in English.

> 8 means the person speaks the words very poorly.
Other ratings would mean that the speaker's performance fell between the extremes of 1 and 8.

There will be a short pause between phrases to permit you to mark your paper.

You will now hear some samples of the range of speakers you are about to rate.

The following instructions were given for Tape B:

You will hear phrases followed by short pauses. For example:

pull ( )
the ( )
stick ( )
back ( )

Put a check in the parenthesis for each word you think is not spoken the way it should be spoken in English. For example, if you think that pull and the are spoken well and stick and back are spoken poorly, put a mark in the parenthesis after stick and back.

Appendix E contains a copy of the answer sheet presented to the judges.

The data for Tape A were tabulated and scale values and Q-values were computed for each passage by the graphic methods suggested by Thurstone and Chave.17


To compute the scale values and Q-values graphically, the frequency distributions of the scale values assigned by the judges were tabulated for each passage. The cumulative frequencies were converted to proportions and plotted on graph paper. The scale values were arranged along the abscissa and the cumulative proportions along the ordinate. A line drawn vertically from the intersection of the 50 per cent point and the plotted curve to the abscissa indicated the scale value. The Q-value was determined by dropping lines from the intersections of the 25 and 75 per cent points.
along the curve and determining the distance between the points of the abscissa.

The results of the judgments for Group A of Tape A are summarized in Table I. The range of pronunciation excellence is divided into eight class intervals from 0.50 to 8.49. The number of samples of nine-second passages, rated from 0.50 to 1.49, is six. The number of samples of nine-second passages in the class interval 7.50 to 8.49 is five. The mean of these samples is represented in the next row. For example, the mean scale value for the six samples of nine-second passages in the class interval 0.50 to 1.49 is 1.04. The mean Q-values of the samples that fell in each of the eight class intervals are found in the bottom row. For example, the mean Q-value for the six nine-second passages that fell in the class interval 0.50 to 1.49 is 0.72.

Tape I (nine-second passages) was constructed from passages selected from Tape A. Two of the passages with the smallest Q-values were selected to represent each of the eight scale values, or sixteen passages in all. Each of the sixteen nine-second passages were connected by a four-second strip of timing tape. The resulting tape will be called Tape I. (nine-second passages).

Table II summarises the samples of nine-second passages selected for Tape I (nine-second passages). The
## Table I

**Scale Values, Mean Scale Values, and Mean Q-Values Assigned to Tape A by 100 College Students (Group A)**

<table>
<thead>
<tr>
<th>Scale Value</th>
<th>Number of Samples</th>
<th>Mean Scale Value</th>
<th>Mean Q-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50-1.49</td>
<td>6</td>
<td>1.04</td>
<td>0.72</td>
</tr>
<tr>
<td>1.50-2.49</td>
<td>7</td>
<td>1.96</td>
<td>1.31</td>
</tr>
<tr>
<td>2.50-3.49</td>
<td>7</td>
<td>2.86</td>
<td>1.45</td>
</tr>
<tr>
<td>3.50-4.49</td>
<td>4</td>
<td>4.02</td>
<td>1.65</td>
</tr>
<tr>
<td>4.50-5.49</td>
<td>11</td>
<td>4.99</td>
<td>1.50</td>
</tr>
<tr>
<td>5.50-6.49</td>
<td>10</td>
<td>5.93</td>
<td>1.46</td>
</tr>
<tr>
<td>6.50-7.49</td>
<td>4</td>
<td>7.05</td>
<td>1.15</td>
</tr>
<tr>
<td>7.50-8.49</td>
<td>5</td>
<td>7.80</td>
<td>1.34</td>
</tr>
</tbody>
</table>
TABLE II
ASSIGNED SCALE VALUES AND MEAN SCALE VALUES FOR TAPE I (NINE-SECOND PASSAGES) FOR THREE GROUPS OF JUDGES

<table>
<thead>
<tr>
<th>Mean Scale Values</th>
<th>Assigned Scale Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group:</td>
</tr>
<tr>
<td>1.2</td>
<td>A -</td>
</tr>
<tr>
<td>0.7</td>
<td>B -</td>
</tr>
<tr>
<td>1.7</td>
<td>C -</td>
</tr>
<tr>
<td>2.4</td>
<td>t</td>
</tr>
<tr>
<td>2.5</td>
<td>t</td>
</tr>
<tr>
<td>3.3</td>
<td>B</td>
</tr>
<tr>
<td>3.5</td>
<td>B</td>
</tr>
<tr>
<td>4.4</td>
<td>A</td>
</tr>
<tr>
<td>5.4</td>
<td>A</td>
</tr>
<tr>
<td>5.4</td>
<td>B</td>
</tr>
<tr>
<td>6.1</td>
<td>B</td>
</tr>
<tr>
<td>6.2</td>
<td>B</td>
</tr>
<tr>
<td>7.1</td>
<td>C</td>
</tr>
<tr>
<td>7.0</td>
<td>C</td>
</tr>
<tr>
<td>8.0</td>
<td>C</td>
</tr>
<tr>
<td>7.9</td>
<td>C</td>
</tr>
</tbody>
</table>

* 100 judges
** 19 judges
*** 33 judges
representative samples of this tape are shown under Mean Scale Values, Group A. The discrete or rounded number corresponding to the mean scale value of the sample appears under "Assigned Scale Values," Group A. The other entries in Table II are discussed below.

Six passages were dubbed at the beginning of Tape I (nine-second passages) as examples of the range of scale values in the test. The mean scale values and the rounded numbers representing the scale values are shown below. The rounded scale values are referred to as "Assigned Scale Values."

<table>
<thead>
<tr>
<th>Assigned Scale Values</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Scale Values</td>
<td>1.1</td>
<td>1.9</td>
<td>3.3</td>
<td>5.0</td>
<td>5.5</td>
<td>7.7</td>
</tr>
</tbody>
</table>

The reliability of Tape I (nine-second passages) was determined by playing the tape to two groups, Group B and Group C. Group B was composed of nineteen graduate students majoring in Speech in the Department of Speech, The Ohio State University. Group C was composed of thirty-three students enrolled in Speech 401. The latter students, thus, were similar in composition to the students in the criterion group. Both groups received the same instructions as Group A.

The mean scale values for Groups A, B, and C, which appear in the first three columns of Table II, were rounded
and the representative (Assigned) values are shown in columns four, five, and six. When the sixteen mean scale values for Groups A and B, A and C, and B and C, shown in columns one, two, and three, were correlated, the Pearson product-moment coefficients of correlation were respectively .88, .96, and .93. When the mean scale values in columns four, five, and six for Groups A and B, A and C, and B and C were correlated, the Pearson product-moment correlations were respectively .97, .95, and .97.

The responses obtained for Groups B and C for Tape I (nine-second passages) were pooled. The means of the responses were computed for each of the sixteen passages for the combined fifty-two subjects. As stated above, Tape I (nine-second passages) was composed of two passages for each of the eight scale values. The means of the responses for the two scale values 1 were arranged next to each other in a column, the means of the responses for scale values 2 were arranged next to each other in a column, and so on, for the eight scale values and sixteen mean responses for the passages. The product-moment correlation when the passages were arranged in this way was .95. When corrected for length by the Spearman-Brown Prophecy formula, the correlation was, \( r = .97 \).

The above computations indicate that this test is reasonably reliable for native speakers of English.
Additional data on the reliability of this test for non-native speakers of English are given in Chapter V.

The Validity of Tape I (nine-second passages). In order to determine whether the passages employed in Tape I (nine-second passages) were testing pronunciation differences, the following null hypothesis was formulated: There is no relationship between the assigned scale values of the passages in Tape I (nine-second passages) and the values of the words contained in the passages when the latter values are based on the number of people in Group A who said the words were not spoken in the way they should be spoken in English.

Tape II (four-word phrases) was also played to Group A. The students were asked to indicate the words that were not spoken the way they should be spoken in English. In this way, a value was determined for each word, based on the number of people who said this word was not spoken the way it should be spoken in English. The value for a given phrase was determined by adding the values for each word in the phrase.

The values for the total words in the passages thus obtained were correlated with the assigned scale values for the passages. The Pearson product-moment correlation was .84, significant at the 1 per cent level of confidence. This rejects the hypothesis that there is no relationship and demonstrates the validity of using these passages to test pronunciation.
The construction of Tape III (pairs). Another test was devised in the following manner to test the ability to judge controlled patterns of speech. The values of Tape II (four-word phrases) described above, were converted to percentage values and arranged in order from the lowest to highest percentage value. The ordered percentage values were then divided into quartiles. Phrases in the first quartile had percentage values of twenty-three or less, that is, these phrases were judged to contain few pronunciation errors. Phrases in the fourth quartile had percentage values of eighty-three or more. From these two groups of phrases, all phrases with percentage values of twenty-three or less representing non-native speakers of English were paired with phrases representing non-native speakers with percentage values of eight-three or more. All phrases with values of twenty-three or less spoken by Americans were paired with phrases with values of eighty-three or more spoken by Americans. This resulted in twenty-four pairs of phrases. Six more pairs of phrases with values near eighty-three and twenty-three were added to bring the total number of pairs to thirty. The phrases of a pair were separated by a one-second pause, and a four-second length of timing tape was inserted between pairs. This resulted in a tape made up of seventeen pairs spoken by Americans and thirteen pairs spoken by international students. This tape will be called Tape III (pairs).
To determine whether scores obtained for Tape III (pairs) would differentiate between native and non-native speakers of English, the tape was played to nineteen graduate students majoring in speech at The Ohio State University. They heard the following instructions:

You will hear pairs of phrases. You are to indicate which of the two phrases is pronounced better. The words in one phrase of the pair are not the same as the words of the other. If you think the first phrase of the pair is better than the second, write 1 on your answer sheet. If you think the second phrase of the pair is pronounced better than the first, write 2 on your answer sheet.

Scores were based on deviations from the values assigned to the phrases by the criterion judges, Group A. Only four of the nineteen judges received perfect scores. The range of errors was 0 - 6. The mean score was 27.94 ± 1.35. No pair was marked incorrectly by more than five judges. Eighteen of the pairs were scored incorrectly by one or more judges.

Tape III (pairs) was also played to twenty-three members of Speech 405 (Speech for International Students). No native speakers of Spanish or Japanese were included among the judges. The range of errors was 3 - 25. The mean score was 19.04 ± 5.31.

A t ratio of 6.74 significant at the 1 per cent level of confidence was found for the difference between the means for the nineteen graduate students and the twenty-three
international students. This ratio indicates that scores obtained from Tape III (pairs) differentiated between native and non-native speakers of English and demonstrates the validity of this test for the determination of pronunciation differences.

Summary of Part I. This part of Chapter III has been devoted to the criteria for selection of three standardized tests and the preparation of three tests designed especially for the purposes of this study. The three standardized tests that were selected were (1) Lado's Test of Aural Comprehension, (2) Templin's Short Test of Speech-sound Discrimination, and (3) a Waco twenty-four word multiple-choice intelligibility test. The three tests especially designed for the purposes of this study were (1) a group test of letter-prediction ability, (2) Tape I (nine-second passages), and (3) Tape III (pairs).

Part II of this chapter is concerned with the criteria for the selection of the experimental subjects who took part in the major part of this study. Part II is divided into two sections. Section A is concerned with the selection of sixty native and non-native speakers of English, and Section B is concerned with the selection of thirty-six international students chosen for the purpose of establishing the effectiveness of Tape I (nine-second passages) as an indirect test of pronunciation.
Part II. Selection of the Experimental Subjects.

Section A — Selection of Sixty Native and Non-native Speakers of English. The criteria for the selection of subjects were these: (1) All the subjects should be approximately equal in intelligence. (2) Native speakers of English should be included to find if the test differentiated between sophisticated and unsophisticated speakers of English. (3) In terms of non-native speakers, backgrounds in the use of English should be approximately equal. (4) The native language of non-native speakers of English, however, should be different.

In order to insure as far as possible that the students were approximately equal in intelligence, college students were chosen as subjects for this study. This meant that the native speakers of English participating in this study had at least finished high school.

In the case of foreign students it was recognized that there would be various degrees of exposure to training in the English language. Lado\(^{18}\) pointed out that previous study of English had little bearing on proficiency in a language. As evidence he stated that 100 random cases were taken from the files of the English Language Institute and

\(^{18}\) Lado, "Measurement in English as a Foreign Language," _op. cit._
were classified into five groups of twenty each according to the student's own statement of the number of hours of previous training in English. The number of hours reported ranged from four to one thousand two hundred. The same 100 cases were classified again according to their scores in aural comprehension at the time of entrance to the Institute.

Out of 100 students, only twenty were in the same group in both number of hours of English studied and aural comprehension; the remaining eighty were in different groups. Nearly half of these (thirty-eight out of eighty) were in groups that were two or more levels apart, and two were at opposite extremes.\(^{19}\)

\(^{19}\) Ibid., pp. 102-3.

Although previous study of English varied considerably among the non-native speakers of English on The Ohio State University and Ohio Wesleyan University campuses, most of them had had little if any experience listening to the speech of native speakers of English before entering the United States. Assuming that continued exposure to English spoken by native speakers of the language would result in improved comprehension and intelligibility, an attempt was made to control this factor. Therefore, subjects were selected who had been in the United States three years or less.
Native speakers of Japanese and Spanish were selected because (1) within each group presumable the subjects would have similar language backgrounds and would encounter similar problems in understanding and speaking English, (2) Spanish is relatively similar to English, and (3) Japanese is quite unlike both Spanish and English. If gross differences do exist between these groups, the tests employed in this study should show these differences.

Sixty subjects were selected for testing for the first part of the study. Twenty were native citizens of the United States and were enrolled in Speech 401 classes at The Ohio State University during the Autumn Quarter, 1953. Eight were males and twelve, females. The mean age was 20.6 years.

Twenty were native speakers of Spanish. One was enrolled at Ohio Wesleyan University. The median residence in the United States for these students was 3.0 months with a range from two to twenty-five months. Nine were females and eleven were males. The mean age was 29.4 years.

The remaining twenty subjects were native speakers of Japanese. Three of the subjects were enrolled at Ohio Wesleyan at the time of the study. The remaining subjects were enrolled at The Ohio State University. The median residence in the United States for these students was 4.0
months. The range was two to thirty-five months. Three were females and seventeen were males. The mean age for this group was 27.5 years.

Although all the Japanese speakers were natives of Japan, the Spanish speakers came from different countries. The native countries of the Spanish-speaking students and the number of students from each country are listed below:

<table>
<thead>
<tr>
<th>Country</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>2</td>
</tr>
<tr>
<td>Bolivia</td>
<td>2</td>
</tr>
<tr>
<td>Chile</td>
<td>2</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>1</td>
</tr>
<tr>
<td>Cuba</td>
<td>4</td>
</tr>
<tr>
<td>Ecuador</td>
<td>2</td>
</tr>
<tr>
<td>Mexico</td>
<td>2</td>
</tr>
<tr>
<td>Panama</td>
<td>1</td>
</tr>
<tr>
<td>Peru</td>
<td>2</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>2</td>
</tr>
</tbody>
</table>

Section B. Selection of Thirty-six Non-native Speakers of English. As a result of the statistical analysis of the data in Chapter IV, a significant correlation was found between the test of aural comprehension and the test of intelligibility. Although Tape I (nine-second passages) correlated significantly with aural comprehension, this measure did not correlate with intelligibility. This result could be interpreted to mean that Tape I (nine-second passages) was not performing the function for which it was designed, i.e., testing pronunciation ability. However, because each speaker was permitted to speak at the sound pressure level of his choice, variations in intensity may have influenced the intelligibility test scores and in
turn the correlation between Tape I (nine-second passages) and intelligibility. To check this another test of intelligibility with better control of the sound pressure level of the speaker's voice was administered to a second group of people.

Statistical analysis of the data in Chapter IV further revealed that Tape I (nine-second passages) differentiated between the native and non-native speakers of English but not between the native speakers of Spanish and Japanese. This result indicated that the test possibly was sensitive enough to differentiate between individuals, but not between groups of subjects that were different in language background.

Because it differentiated between the native and non-native speakers, it provided evidence for the validity of this test as an indirect measure of intelligibility. In any event, consideration was given to the fact that Tape I (nine-second passages) might be useful as an index of the intelligibility of international students.

A second group of thirty-six subjects was tested. None of this group spoke English as a native language. The native language and the number of subjects speaking the language are given below.
Thirty-one of the subjects were males. Five were females. The mean age was 28.4 years. The range was twenty—fifty-six years with a median of 27.0 years. Mean residence in the United States was 12.9 months. The range was two—forty-three months with a median of 5.0 months.

The test of aural comprehension and Tape I (nine-second passages) were administered to each subject in the same manner as above. In addition, each subject was required to read one list of the twenty-two lists for the Waco twenty-four word write-down intelligibility test.20

---


This test was selected to replace the twenty-four word multiple-choice test that was used to test intelligibility for the sixty native and non-native speakers of English because the techniques for saying the words for this test are
more amenable to control of rate and sound pressure level. The responses of the subjects were recorded as they read words on a Type PT 6-P3 Magnecorder at seven and one-half inches per second. In this instance each subject was asked to monitor the intensity of his voice by keeping the indicator for the VU meter on the Magnecorder approximately on a line drawn on the face of the meter. Each subject was given a chance to practice visual monitoring of his voice by reading a different list from the one that would be recorded. Each word was introduced by "number (the ordinal number) is." The time between phrases was measured by a sweep-second timer; five seconds elapsed between each phrase. The lists were played to three groups of twelve American students each. These students were enrolled in Speech 401, Autumn Quarter, 1952. The students were given an answer sheet and were asked to write the words as they heard them.

The tapes were played at approximately 65 ± 5 db measured on a General Radio Company Sound Level Meter, Type 759-B (C-scale) measured six feet from the output speaker. The tapes were reproduced on a Record-O-Fone recorder reproducer connected to an Altec Speaker as in the procedure for Section A. An electric fan was turned on while the tapes were being played to introduce noise into the listening environment.
Part III of this chapter is devoted to the procedures followed in testing the two groups of sixty and thirty-six experimental subjects.

Part III. The Procedures Followed in Testing.

Six tests were administered to twenty native speakers of English, twenty native speakers of Spanish, and twenty native speakers of Japanese. The tests were (1) Lado's "Test of Aural Comprehension" (2) Templin's "Short Test of Sound Discrimination" (3) Form B of the Waco Multiple-Choice Intelligibility Test (4) A test of letter-prediction ability (5) Tape I (nine-second passages) (6) Tape III (pairs).

No attempt was made to separate the subjects as to language background during the testing sessions. As scheduling was a problem, the number of people in each session varied from one person to as many as ten.

Each person was requested to give his name, native country, age to the nearest birthday, sex, and, in the case of the non-native speakers, the date of entry into the United States. Each person was provided with a set of directions and answer sheets. The answer sheets were the same for the
"Short Test of Sound Discrimination," Tape I (nine-second passages), and Tape III (pairs). Although the number of items in these tests varied from sixteen in the case of Tape I (nine-second passages) to seventy in the case of the "Short Test of Sound Discrimination," the subjects were not told the number of items they would hear.

Tape I (nine-second passages), Tape II (pairs), Lado's "Test of Aural Comprehension," and Templin's "Short Test of Sound Discrimination" were played at the same volume level in all situations. The tapes were played on a Bell Record-O-Fone tape recorder reproducer Model RT-65B connected to an Altec Speaker Model 400B mounted in a baffle.

The order of the presentation of tapes was changed in every testing situation. This was done by assigning each tape a number, placing this number on a card, shuffling the cards for each session, and playing the tapes in the order indicated by the cards. As there was no time limit for the test of letter-prediction ability, this measure was given last. After the subjects had begun work on the letter-prediction test they were taken aside, one at a time, to make the recording for intelligibility testing.

Summary of Chapter III. This chapter was concerned with the methods and procedures for the selection and preparation of the materials, selection of the experimental
subjects, and testing the experimental subjects. Part I of this chapter was concerned with a statement of the criteria for the selection and preparation of six tests: (1) Lado's Test of Aural Comprehension, (2) Templin's Short Test of Speech-sound Discrimination, (3) a Waco twenty-four word multiple-choice intelligibility test, (4) a group test of letter-prediction ability, (5) Tape I (nine-second passages) and (6) Tape III (pairs). Part II states the criteria for the selection of two groups of experimental subjects and Part III described the controls employed relative to the testing procedures.
CHAPTER IV

RESULTS AND DISCUSSION

PART I. Results Based on Scores Obtained for Sixty Subjects.

The Computation of the Scores for the Six Tests. Six tests were administered to twenty English-speaking students, twenty Spanish-speaking students, and twenty Japanese-speaking students: (1) a test of aural comprehension, (2) a test of sound discrimination, (3) and (4) two measures designed to test preferences for controlled speech patterns, (5) a test of intelligibility, and (6) a measure designed to test the ability to predict the letters of English writing.

Aural Comprehension. The scores for the subjects for the Test of Aural Comprehension were obtained by recording the stimulus sentences on magnetic tape and playing them to the subjects. The subjects heard sixty stimulus sentences and responded to standardized multiple-choice items. The number of items answered incorrectly by a subject was subtracted from sixty (the total possible number of correct answers) to obtain his score.

Speech-sound Discrimination. The stimulus material for the Short Test of Sound Discrimination was recorded on magnetic tape and played to the subjects. The subjects noted whether the pairs of syllables were the "same" or "different"
The responses of a subject were compared with the stimulus material. The number of errors in judgment was subtracted from seventy (the total number of pairs of syllables) to obtain the score for a subject.

**Tape I** (*nine-second passages*). The two measures designed to test preferences for controlled speech patterns were Tape I (*nine-second passages*) and Tape III (*pairs*). Tape I (*nine-second passages*) consisted of sixteen passages. Each passage was nine seconds in length and was spoken by a native or non-native speaker of English. The passages previously had been rated on an eight-point scale in terms of pronunciation excellence by 100 native speakers of English. A scale value of 1 meant that the words in the passages were spoken the way they should be spoken in English. These passages were recorded and played to the experimental subjects who were asked to rate these passages on the eight-point scale. The sum of the deviations from the ratings of the criterion judges was taken as the error scores for the experimental subjects. The direction of the deviation was ignored. Thus, a score of zero on this test would indicate agreement with the ratings of the 100 native speakers of English.

**Tape III** (*pairs*). This test consisted of thirty pairs of four-word phrases spoken by native and non-native speakers of English. The words of each pair of phrases had
been rated by 100 native speakers of English in terms of pronunciation excellence. One member of each pair had been rated as better than the other in terms of pronunciation by the criterion judges. These phrases were recorded on magnetic tape and played to the experimental subjects. Error scores for each subject were obtained by noting the deviations from the ratings of the criterion judges. The error scores were subtracted from thirty (the total number of pairs) to obtain the score for each subject.

**Intelligibility.** One of the twenty-four word lists for Form B of the Waco multiple-choice intelligibility test was read by each subject. These lists were recorded and played to groups of native speakers of English; each group consisted of twenty college students. These students heard the words spoken by the speakers and selected the words from multiple-choice items. Per cent intelligibility was determined for each speaker.

**Letter Prediction.** To obtain the scores for the measure designed to test the ability to predict the letters of English, each subject was presented a set of six answer sheets. Each answer sheet consisted of fifteen columns of twenty-seven symbols each and was coded by means of a chemical solution with a message fifteen letters in length. Each subject predicted the letters by marking the symbols in each column with green ink. When the correct
letter was found, the green ink turned brown. The number of trials for each letter was entered at the top of the column. Three scores were obtained for each subject: (1) the total number of trials necessary to predict the ninety letters in the six passages (no more than ten trials were counted for each letter), (2) the total number of letters predicted correctly on the first trial, and (3) the total number of letters predicted correctly after ten or more previous trials.

The raw scores for the subjects for the six tests are enumerated in Appendix G.

A summary of the statistical methods employed in analyzing the obtained data is summarized in Table III.

Graphic Presentation of the Scores. Figures 1 through 6 graphically show the distribution of the scores for the three groups of subjects: native speakers of English, Spanish, and Japanese. The ordinates of the Figures represent the proportion of twenty students that obtained scores on the various tests. The circles represent the distribution of the obtained scores. The curves were faired through the circles. Figure 1 shows the cumulative distribution of raw scores for the test of aural comprehension. The possible range of scores for this measure is 0-60 errors. The range of scores for the pooled groups is 26-60. The range of scores for the Spanish-speaking group is 26-60 with a mean of 46.6 ± 7.86. For the Japanese-speaking group the range is 37-57. The mean is 49.2 ± 6.06. The scores for the English-speaking group cover the smallest
<table>
<thead>
<tr>
<th>TESTS</th>
<th>STATISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tests of relationship for the Japanese-, Spanish-, and English-speaking groups for the scores obtained for (1) a test of aural comprehension, (2) a test of sound discrimination, (3) and (4) two tests of preferences for controlled speech patterns, (5) a test of intelligibility, (6) a test of letter prediction.</td>
<td>Product-moment correlation</td>
</tr>
<tr>
<td>2. Tests of relationship for the pooled groups of Japanese- and Spanish-speaking groups for the scores mentioned above.</td>
<td>Product-moment correlation</td>
</tr>
<tr>
<td>3. A comparison of mean scores of each of the Japanese-, Spanish-, and English-speaking groups for each of the six measures mentioned above.</td>
<td>t ratio for differences between means</td>
</tr>
<tr>
<td>4. Tests of relationship for the pooled Japanese- and Spanish-speaking groups among the scores mentioned above.</td>
<td>Product-moment correlation</td>
</tr>
<tr>
<td>5. A comparison of the Spanish-, Japanese-, and English-speaking groups for: (1) the number of letters predicted correctly on the first trial, and (2) the number of letters guessed after ten or more previous trials.</td>
<td>F ratio</td>
</tr>
</tbody>
</table>
Figure 1

Plot of Cumulative Scores Obtained for 20 Native Speakers of Spanish, 20 Native Speakers of Japanese, and 20 Native Speakers of English for the Test of Aural Comprehension
range (seven scale points). The range is 53-60 with a mean of 58.5 ± 1.77. Eight of the Americans scored 100 per cent on this test, that is 60. One Spanish-speaking student scored 100 per cent.

Figure 2 shows the cumulative distribution of scores for sound discrimination. The possible range of scores for this measure is 0-70 errors. The range of scores for the combined groups is 40-66. The range of scores for the Spanish group is 43-63 with a mean of 56.1 ± 5.38. The scores for the Japanese group is 40-64 with a mean of 57.2 ± 5.24. The range for the English-speaking subjects is 55-66 with a mean of 60 ± 2.96.

Figure 3 represents the cumulative distribution of scores for Tape I (nine-second passages). The possible range of scores is 88-0. The range for the combined groups is 12-144. The range of scores for the Spanish group is 12-42 with a mean of 28.2 ± 8.64. The range for the Japanese group is 17-44 with a mean of 29.6 ± 7.27. The range for the English-speaking group is 12-30 with a mean of 18.4 ± 5.26.

Figure 4 represents the cumulative range of scores for Tape III (pairs). Although the possible range of scores is 0-30, the obtained scores for the pooled groups fall within a range of 15-25. The mean for the Japanese group is 19.1, for the Spanish group 19.55, and for the English-
Figure 2

Plot of Cumulative Scores Obtained for 20 Native Speakers of Spanish, 20 Native Speakers of Japanese, and 20 Native Speakers of English for the Test of Sound Discrimination
Figure 3

Plot of Cumulative Scores Obtained for 20 Native Speakers of Spanish, 20 Native Speakers of Japanese, and 20 Native Speakers of English for Tape I (nine-second passages)
Figure 4

Plot of Cumulative Scores Obtained for 20 Native Speakers of Spanish, 20 Native Speakers of Japanese, and 20 Native Speakers of English for Tape III (pairs)
speaking group 21.55. The standard deviation for the Japanese group was 2.45, for the Spanish group 2.64, and for the English-speaking group 1.58.

Figure 5 represents the cumulative distribution of per cent intelligibility scores obtained for the Waco multiple-choice intelligibility test. The scores are based on the judgments of twenty judges for each speaker. The range of scores for the pooled groups is 30.2-88.8. The range for the English-speaking group is 35.8-88.8 with a mean of 66.6 ± 11.3. The range for the Japanese-speaking group is 38.5-68.0 with a mean of 50.0 ± 10.0. The range for the Spanish-speaking group is 30.2-78.4 with a mean of 54.0 ± 11.62.

Figure 6 represents the cumulative distribution for the obtained scores for the letter prediction test. The possible range of scores is 0-900. The range of scores obtained for the pooled groups is 247-504 which happens to be the range of scores for the Spanish-speaking group. The Spanish group had a mean of 354.2 ± 62.91. The range of scores for the Japanese group is 261-422 with a mean of 322.7 ± 37.69. The range of scores for the English-speaking group is 250-362 with a mean of 300.35 ± 26.31.

All the mean scores for the three groups were different from chance or guessed scores at the 1 per cent level of confidence (see Table IV). For nineteen degrees of
Plot of Cumulative Scores Obtained for 20 Native Speakers of Spanish, 20 Native Speakers of Japanese, and 20 Native Speakers of English for the Waco Multiple-choice Intelligibility Test
Figure 6

Plot of Cumulative Scores Obtained for 20 Native Speakers of Spanish, 20 Native Speakers of Japanese, and 20 Native Speakers of English for the Test of Letter Prediction
### TABLE IV

**t** RATIOS FOR THE DIFFERENCE BETWEEN OBTAINED MEANS ($M_g$) AND GUESSED MEANS ($M_o$) FOR THREE GROUPS OF SUBJECTS FOR SIX TESTS

<table>
<thead>
<tr>
<th></th>
<th>$M_g$</th>
<th>$M_o$</th>
<th>S.D.$M_o$</th>
<th>S.E.M.$M_o$</th>
<th>$t^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aural Comprehension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>20</td>
<td>46.8</td>
<td>7.86</td>
<td>1.74</td>
<td>15.40</td>
</tr>
<tr>
<td>Japanese</td>
<td>20</td>
<td>49.2</td>
<td>6.06</td>
<td>1.37</td>
<td>22.10</td>
</tr>
<tr>
<td>English</td>
<td>20</td>
<td>58.5</td>
<td>1.77</td>
<td>.40</td>
<td>96.20</td>
</tr>
<tr>
<td><strong>Sound Discrimination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>35</td>
<td>56.1</td>
<td>5.38</td>
<td>1.22</td>
<td>17.29</td>
</tr>
<tr>
<td>Japanese</td>
<td>35</td>
<td>57.2</td>
<td>5.24</td>
<td>1.19</td>
<td>18.68</td>
</tr>
<tr>
<td>English</td>
<td>35</td>
<td>60.0</td>
<td>2.96</td>
<td>.67</td>
<td>37.32</td>
</tr>
<tr>
<td><strong>Tape I (nine-second passages)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>44</td>
<td>28.2</td>
<td>8.64</td>
<td>1.93</td>
<td>8.20</td>
</tr>
<tr>
<td>Japanese</td>
<td>44</td>
<td>29.6</td>
<td>7.27</td>
<td>1.63</td>
<td>8.76</td>
</tr>
<tr>
<td>English</td>
<td>44</td>
<td>18.4</td>
<td>5.26</td>
<td>1.18</td>
<td>21.65</td>
</tr>
<tr>
<td><strong>Tape III (pairs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>15</td>
<td>19.5</td>
<td>2.64</td>
<td>.60</td>
<td>7.59</td>
</tr>
<tr>
<td>Japanese</td>
<td>15</td>
<td>19.1</td>
<td>2.45</td>
<td>.55</td>
<td>7.46</td>
</tr>
<tr>
<td>English</td>
<td>15</td>
<td>21.5</td>
<td>1.58</td>
<td>.36</td>
<td>18.08</td>
</tr>
<tr>
<td><strong>Intelligibility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>50</td>
<td>54.0</td>
<td>11.62</td>
<td>2.59</td>
<td>7.99</td>
</tr>
<tr>
<td>Japanese</td>
<td>50</td>
<td>50.0</td>
<td>10.00</td>
<td>2.25</td>
<td>9.19</td>
</tr>
<tr>
<td>English</td>
<td>50</td>
<td>66.6</td>
<td>1.30</td>
<td>2.52</td>
<td>14.29</td>
</tr>
<tr>
<td><strong>Letter Prediction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>450</td>
<td>354.2</td>
<td>62.91</td>
<td>14.07</td>
<td>6.77</td>
</tr>
<tr>
<td>Japanese</td>
<td>450</td>
<td>322.7</td>
<td>37.69</td>
<td>8.43</td>
<td>15.10</td>
</tr>
<tr>
<td>English</td>
<td>450</td>
<td>300.3</td>
<td>26.31</td>
<td>5.88</td>
<td>25.41</td>
</tr>
</tbody>
</table>

* For 19 degrees of freedom, $t = 2.86$ is significant at the 1 per cent level of confidence.
freedom, a \( t \) of 2.86 is significant at the 1 per cent level of confidence.

The purpose of this study was to test four null hypotheses concerning English as a foreign language. The first two of these hypotheses were:

**Hypothesis I**

**THERE IS NO DIFFERENCE BETWEEN NATIVE AND NON-NATIVE SPEAKERS OF ENGLISH (A) IN THE ABILITY TO PREDICT THE LETTERS OF ENGLISH, (B) IN THE ABILITY TO COMPREHEND ENGLISH, (C) IN THE ABILITY TO DIFFERENTIATE BETWEEN THE SOUNDS OF ENGLISH, (D) IN INTELLIGIBILITY WHEN SPEAKING ENGLISH, AND (E) IN THE ABILITY TO JUDGE CONTROLLED PATTERNS OF SPEECH.**

**Hypothesis II**

**THERE IS NO DIFFERENCE BETWEEN NATIVE SPEAKERS OF SPANISH AND JAPANESE IN THE VARIABLES STATED IN HYPOTHESIS I.**

**Tests of the Differences of the Means.** In order to test the two hypotheses, the means, standard deviations, and standard error of the means were computed (see Table V). Table VI shows the \( t \) ratios for the tests of the significance of the difference among the means for the three groups. With the single exception of the scores for sound discrimination, the means for the English-speaking groups were significantly different from the means for the Japanese- and Spanish-
<table>
<thead>
<tr>
<th>Measure</th>
<th>Spanish*</th>
<th>Japanese*</th>
<th>English*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M S.D. S.E.M.</td>
<td>M S.D. S.E.M.</td>
<td>M S.D. S.E.M.</td>
</tr>
<tr>
<td>Aural Comprehension</td>
<td>46.80 7.86 1.76</td>
<td>49.20 6.06 1.35</td>
<td>58.50 1.77 0.38</td>
</tr>
<tr>
<td>Sound Discrimination</td>
<td>56.10 5.38 1.20</td>
<td>57.20 5.24 1.17</td>
<td>60.70 2.96 0.66</td>
</tr>
<tr>
<td>Tape I (nine-second passages)</td>
<td>28.20 8.64 1.93</td>
<td>29.60 7.27 1.63</td>
<td>18.40 5.26 1.18</td>
</tr>
<tr>
<td>Tape III (pairs)</td>
<td>19.55 2.64 0.59</td>
<td>19.10 2.45 0.55</td>
<td>21.55 1.58 0.35</td>
</tr>
<tr>
<td>Intelligibility</td>
<td>54.01 11.55 2.59</td>
<td>50.03 10.06 2.25</td>
<td>66.60 11.30 2.52</td>
</tr>
<tr>
<td>Letter Prediction</td>
<td>354.20 62.91 14.07</td>
<td>322.70 37.69 8.43</td>
<td>300.35 26.31 5.88</td>
</tr>
</tbody>
</table>

* N = 20
### TABLE VI

**t** RATIOS FOR TESTS OF SIGNIFICANT DIFFERENCES AMONG THE MEANS OF THREE GROUPS FOR SIX MEASURES

<table>
<thead>
<tr>
<th>Test</th>
<th>American and Japanese Subjects</th>
<th>American and Spanish Subjects</th>
<th>Japanese and Spanish Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aural Comprehension</td>
<td>6.65**</td>
<td>6.50**</td>
<td>1.08</td>
</tr>
<tr>
<td>Sound Discrimination</td>
<td>1.33</td>
<td>3.35**</td>
<td>.67</td>
</tr>
<tr>
<td>Tape I (nine-second passages)</td>
<td>5.54**</td>
<td>4.03**</td>
<td>.56</td>
</tr>
<tr>
<td>Tape III (pairs)</td>
<td>3.75**</td>
<td>2.94**</td>
<td>.50</td>
</tr>
<tr>
<td>Intelligibility</td>
<td>5.43**</td>
<td>3.60**</td>
<td>1.14</td>
</tr>
<tr>
<td>Letter Prediction</td>
<td>2.18*</td>
<td>3.53**</td>
<td>1.92</td>
</tr>
</tbody>
</table>

* 5 per cent level of confidence  
** 1 per cent level of confidence  

N = 20 for each group
speaking groups at the 5 per cent level of confidence.\(^1\)

\(^1\) The \(t\) ratios were obtained by using the following formula for a \(t\) ratio for a difference between means:

\[
t = \frac{M_1 - M_2}{\sqrt{\frac{2}{\sigma_{M_1}^2 + \sigma_{M_2}^2}}} \text{ where } \frac{dM}{\sigma_{M_1}} = \text{S.E. of the mean of the first distribution and } \frac{dM}{\sigma_{M_2}} = \text{S.E. of the mean of the second distribution.}
\]


In the case of the scores obtained for speech-sound discrimination, the English-speaking group and the Spanish-speaking group had significantly different means at the 1 per cent level of confidence. The Japanese group was not significantly different from the English or the Spanish groups at the 5 per cent level of confidence.

The data for letter prediction summarized in Table V was based on the total trials for each letter for the subjects, counting no more than ten trials for each letter. Another analysis was made after the number of letters predicted correctly on the first trial for each subject was determined. Table VII summarises the means, standard deviations, and standard error of the means for the data for letter prediction based on the number of letters predicted correctly on the first trial. These figures are based on a total of ninety letters.
An F ratio of 8.42 significant at the 1 per cent level of confidence was obtained for the scores of the three groups. This result indicated significant differences among the groups (see Table VIII). Tests of significant differences between the means again showed that the obtained means for the native speakers and the non-native speakers were significantly different. Again, the means for the Spanish and Japanese groups were not significantly different. The t ratios for the groups were:

<table>
<thead>
<tr>
<th></th>
<th>English vs. Spanish</th>
<th>English vs. Japanese</th>
<th>Spanish vs. Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td>t ratio</td>
<td>4.15</td>
<td>3.34</td>
<td>1.23</td>
</tr>
<tr>
<td>Significant at</td>
<td>1 per cent level</td>
<td>1 per cent level</td>
<td>Not significant</td>
</tr>
<tr>
<td>the</td>
<td>of confidence</td>
<td>of confidence</td>
<td>at the 5 per cent</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>level of confidence</td>
</tr>
</tbody>
</table>
TABLE VIII

SUMMARY OF ANALYSIS OF VARIANCE OF SCORES BASED ON THE NUMBER OF LETTERS PREDICTED CORRECTLY ON THE FIRST TRIAL BY NATIVE SPEAKERS OF SPANISH, JAPANESE, AND ENGLISH

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Variance</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>2</td>
<td>711.45</td>
<td>355.72</td>
<td>8.42*</td>
</tr>
<tr>
<td>Within groups</td>
<td>57</td>
<td>2406.89</td>
<td>42.23</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>3118.34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the 1 percent level of confidence.
The number of letters predicted correctly after ten previous trials was computed for each individual. The means, standard deviations, and standard error of the means for this analysis are listed in Table IX.

TABLE IX

THE MEANS, STANDARD DEVIATIONS, AND STANDARD ERROR OF THE MEANS OBTAINED FROM THE DATA FOR LETTER PREDICTION BASED ON THE NUMBER OF LETTERS PREDICTED CORRECTLY AFTER TEN PREVIOUS TRIALS FOR NINETY LETTERS

<table>
<thead>
<tr>
<th></th>
<th>English*</th>
<th>Spanish*</th>
<th>Japanese*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>11.85</td>
<td>16.05</td>
<td>13.90</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.68</td>
<td>6.78</td>
<td>4.79</td>
</tr>
<tr>
<td>Standard Error of the Mean</td>
<td>.82</td>
<td>1.52</td>
<td>1.07</td>
</tr>
<tr>
<td>N = 20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An F ratio of 3.15 was found, significant at the 5 per cent level of confidence. This F ratio indicated significant differences among the three groups (See Table X). The t ratios for the groups were:

<table>
<thead>
<tr>
<th></th>
<th>English vs. Spanish</th>
<th>English vs. Japanese</th>
<th>Spanish vs. Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.44</td>
<td>1.52</td>
<td>1.56</td>
</tr>
</tbody>
</table>

Significant at the 5 per cent level of confidence  Not significant at the 5 per cent level of confidence  Not significant at the 5 per cent level of confidence
**TABLE X**

**SUMMARY OF ANALYSIS OF VARIANCE OF SCORES BASED ON THE NUMBER OF LETTERS PREDICTED CORRECTLY AFTER TEN PREVIOUS TRIALS BY NATIVE SPEAKERS OF SPANISH, JAPANESE, AND ENGLISH**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Variance</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>2</td>
<td>176.44</td>
<td>88.22</td>
<td>3.15*</td>
</tr>
<tr>
<td>Within groups</td>
<td>57</td>
<td>1595.30</td>
<td>27.98</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>1771.74</td>
<td>* Significant at the 5 per cent level of confidence.</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the 5 per cent level of confidence.
Only the means for the English-speaking subjects and the Spanish-speaking subjects were significantly different at the 5 per cent level of confidence.

**Graphic Presentation of the Scores for the Test of Letter-prediction Ability.** Figures 7 through 12 represent the responses of the three groups for letter prediction. Each curve is based on the mean number of trials for twenty subjects necessary to predict a letter. No more than ten trials per letter per subject were counted.

The curves in Figure 13 represent the mean number of trials needed to predict the letters when the responses for the six phrases were pooled for the native and non-native speaker groups. Each curve was smoothed by averaging the mean responses for every three letters. The curve for the non-native speakers is above the curve for the native speakers. In other words, the non-native speakers averaged more trials per letter than the native speakers. The mean number of guesses necessary to predict a letter was 3.5 for the native speakers of English and 3.8 for the non-native speakers. It will be noted that there is a tendency on the part of the groups to require fewer trials to predict letters as more and more letters are successfully predicted. Up to six letters, with all phrases pooled, the non-native and native groups required about the same number of guesses per letter. By letter nine and after, the difference between
Figure 7:
Mean number of trials for the prediction of the letters "in" and "around."
Figure 5

Keep number of trials for the prediction of the better. "Care at a distance."
Figure 13

Mean Number of Trials Per Sequence of Three Letters (phrases pooled)
the two groups is significant at the 1 per cent level of confidence. The rise in the curve from the twelfth letter to the fifteenth is viewed as an artifact due to the fact that the last word in every phrase was incomplete. Although the directions for the subjects stated that the last word might or might not be complete, there may have been a tendency for the subjects to attempt letters for words that would fit the remaining spaces as in a cross-word puzzle.

Figure 14 represents the mean number of letters guessed correctly on the first trial and the mean number of letters guessed correctly after ten previous trials. As shown in the graph, the native speakers of English predicted more letters on the first trial and less letters after ten previous trials than did either of the non-native speaker groups.

Null Hypothesis I was: THERE IS NO DIFFERENCE AMONG NATIVE SPEAKERS OF ENGLISH, SPANISH, AND JAPANESE IN COMPREHENSION OF ENGLISH, INTELLIGIBILITY, SPEECH-SOUND DISCRIMINATION, PREFERENCES FOR CONTROLLED SPEECH PATTERNS, AND THE ABILITY TO PREDICT THE LETTERS OF ENGLISH.

Null Hypothesis II was: THERE IS NO DIFFERENCE BETWEEN NATIVE SPEAKERS OF SPANISH AND JAPANESE IN THE VARIABLES MENTIONED IN HYPOTHESIS I.
Mean Number of 90 Letters Predicted Correctly on the First Trial and After Ten Previous Trials for Three Groups of Experimental Subjects.
Null hypothesis I is rejected with confidence as regards the difference between the native and non-native speakers of English. Null hypothesis II was not rejected inasmuch as no significant differences were found between the Spanish and Japanese groups. In the case of the test of sound discrimination the difference between the means for the English- and Japanese-speaking groups was not significantly different at the 5 per cent level of confidence. In the case of the test of letter prediction, the differences between the means based on the means obtained for the number of trials predicted correctly after ten or more previous trials were not significantly different at the 5 per cent level of confidence for the English- and Japanese-speaking groups and the Spanish- and Japanese-speaking groups.

Hypothesis III

There is no correlation among scores obtained for a test of aural comprehension, a test of speech-sound discrimination, two measures designed to test the ability to judge controlled patterns of speech, a test of intelligibility, and a test of the ability to predict the letters of English when the tests are administered to native speakers of English, Japanese, and Spanish. This series of hypotheses was tested in the following manner.

The scores obtained for each measure were correlated with every other measure within each group of subjects.
Fifteen product-moment correlations were obtained for each group of subjects in this manner. The correlations for the Spanish-, the Japanese-, and the English-speaking group are shown in Table XI. All the signs for the significant correlations in Table XI are in the direction that would be expected if the tests were measuring what they were designed to measure.

Significant correlations for the Spanish-speaking subjects were found between the test of aural comprehension and Tape I (nine-second passages). The correlation was, $r = -0.50$, significant at the 5 per cent level of confidence. Aural Comprehension and Tape III (pairs) was, $r = 0.59$, significant at the 1 per cent level of confidence. Aural comprehension and letter prediction ability was significant at the 5 per cent level of confidence. The correlation was, $r = -0.49$.

Significant correlations were found for the Japanese students between aural comprehension and letter prediction ($-0.48$) and between sound discrimination and Tape I (nine-second passages) ($-0.45$). Both were significant at the 5 per cent level of confidence. No significant correlation was found among the tests for the native speakers of English.

When the scores for the Japanese- and Spanish-speaking groups were pooled, four significant correlations were found (see Table XII). Aural Comprehension correlated with
<table>
<thead>
<tr>
<th></th>
<th>B.</th>
<th>C.</th>
<th>D.</th>
<th>E.</th>
<th>F.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Aural Comprehension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>.05</td>
<td>-.50*</td>
<td>.59**</td>
<td>.40</td>
<td>-.49*</td>
</tr>
<tr>
<td>Japanese</td>
<td>-.13</td>
<td>-.42</td>
<td>-.01</td>
<td>.41</td>
<td>-.48#</td>
</tr>
<tr>
<td>English</td>
<td>-.20</td>
<td>-.29</td>
<td>.01</td>
<td>.27</td>
<td>.30</td>
</tr>
<tr>
<td><strong>B. Sound Discrimination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td>-.01</td>
<td>-.25</td>
<td>-.09</td>
<td>-.20</td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td>-.45*</td>
<td>.19</td>
<td>.27</td>
<td>-.03</td>
</tr>
<tr>
<td>English</td>
<td></td>
<td>.25</td>
<td>.10</td>
<td>-.08</td>
<td>-.07</td>
</tr>
<tr>
<td><strong>C. Tape I (nine-second passages)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td>-.38</td>
<td>.22</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td>-.04</td>
<td>-.02</td>
<td>.29</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td></td>
<td>-.38</td>
<td>.07</td>
<td>-.25</td>
<td></td>
</tr>
<tr>
<td><strong>D. Tape III (pairs)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td>.12</td>
<td>-.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td>-.07</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td></td>
<td>.05</td>
<td>.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>E. Intelligibility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td></td>
<td></td>
<td>-.21</td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td></td>
<td></td>
<td></td>
<td>-.31</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td>.28</td>
<td></td>
</tr>
<tr>
<td><strong>F. Letter Prediction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**## 1 per cent level of confidence**

*** 5 per cent level of confidence**
<table>
<thead>
<tr>
<th></th>
<th>A. Aural Comprehension</th>
<th>B. Sound Discrimination</th>
<th>C. Tape I (nine-second passages)</th>
<th>D. Tape III (pairs)</th>
<th>E. Intelligibility</th>
<th>F. Letter Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>-.01</td>
<td>-.05</td>
<td>-.22</td>
<td>.002</td>
<td>-18</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-.47**</td>
<td>-.05</td>
<td>.12</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>.32*</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>.36*</td>
<td>-.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>-.51**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Significant at the 1 per cent level of confidence
* Significant at the 5 per cent level of confidence
Tape I (nine-second passages) (−.47) and with letter prediction (−.51) at the 1 per cent level of confidence. Tape III (pairs) (.32) and Intelligibility (.36) correlated with aural comprehension at the 5 per cent level of confidence. Note that speech-sound discrimination correlated significantly only with Tape I (nine-second passages) for the Japanese group. The range of scores for the test of sound-discrimination was decidedly limited, and as shown, the mean scores for the test of speech-sound discrimination differentiated between only the English and Spanish groups.

Hypothesis III was: **THERE IS NO CORRELATION AMONG SCORES OBTAINED FOR A TEST OF AURAL COMPREHENSION, A TEST OF SPEECH-SOUND DISCRIMINATION, TWO MEASURES DESIGNED TO TEST THE ABILITY TO JUDGE CONTROLLED PATTERNS OF SPEECH, A TEST OF INTELLIGIBILITY, AND A TEST OF THE ABILITY TO PREDICT THE LETTERS OF ENGLISH WHEN THE TESTS ARE ADMINISTERED TO NATIVE SPEAKERS OF ENGLISH, JAPANESE, AND SPANISH.**

Because the correlations between aural comprehension and Tape I (nine-second passages), aural comprehension and letter prediction, and speech-sound discrimination and Tape I (nine-second passages) were correlated significantly, the hypothesis that there is no relationship is rejected insofar as concerns five of the tests. The null hypothesis is not rejected that the scores obtained for the test of
Intelligibility are related to the scores for the other five tests. The hypothesis is rejected that the scores obtained for the test of speech-sound discrimination are related to the scores obtained for the other tests only insofar as the relationship between the test of speech-sound discrimination and Tape I (nine-second passages) for the Japanese group is concerned.

PART II. Results Based on Scores Obtained for Thirty-six International Students.

Hypothesis IV.

Judgments of controlled speech patterns are not reliable or valid estimates of intelligibility.

As shown above, twenty native speakers of English and forty non-native speakers of English were administered Tape I (nine-second passages), Tape III (pairs), and a Waco multiple-choice intelligibility test. When the Spanish- and Japanese-speaking groups were pooled no significant correlations were found among Tapes I and III and intelligibility.

As was shown in Figure 4, the range of scores for Tape III (pairs) was decidedly limited; the range was only ten errors for the combined groups. This seems to indicate that this measure was not sensitive enough to differentiate among the subjects.
Intensity fluctuations might have adversely affected the intelligibility scores and in turn influenced the correlation between intelligibility and Tape I (nine-second passages) because sound pressure level was not controlled during the recording of the intelligibility test.

A second group of thirty-six international students was tested. None of this group was a native speaker of English. These subjects were administered Lado's Test of Aural Comprehension and Tape I (nine-second passages). In addition, each subject read one list of the Waco twenty-four word write-down intelligibility test. Figures 15 through 17 show the distribution of scores for the three measures. Appendix H contains the raw scores for the three tests.

As is shown in Figure 15, the range of scores for aural comprehension is 30-60 with a mean of 50.53 \pm 7.36. Figure 16 shows the distribution of error scores for Tape I (nine-second passages). The range is 12-43 with a mean of 28.28 \pm 7.58. The percent intelligibility scores for the Waco twenty-four word write-down test are shown in Figure 17. The scores are arranged in class intervals of five scale points each. The range of scores is 29-152 with a mean of 83.06 \pm 32.14.
Figure 15

Distribution of scores for 36 Non-native Speakers of English for Test of Aural Comprehension
Figure 16

Distribution of Scores for 36 Non-native Speakers of English for Tape I (nine-second passages)
Possible Range of Scores 0-100

Figure 17

Distribution of scores for 30 locative Speakers of English for the Crit-Com Test of Intelligibility

126
The product-moment correlations among the measures are:

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Aural Comprehension</td>
<td>-.58**</td>
<td>.42**</td>
</tr>
<tr>
<td>B. Tape I (nine-second passages)</td>
<td></td>
<td>-.33*</td>
</tr>
<tr>
<td>C. Intelligibility</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Significant at the 1 per cent level of confidence
* Significant at the 5 per cent level of confidence

The correlation for aural comprehension and Tape I (nine-second passages) is -.58, significant at the 1 per cent level of confidence. This correlation compares favorably with the $r$ of -.47 obtained for the forty Japanese- and Spanish-speaking students.

The correlation obtained for aural comprehension and intelligibility was .42 again comparing favorably with the $r$ of .36 for the Japanese and Spanish groups.

The correlation for Tape I (nine-second passages) and intelligibility was, $r = -.33$, significant at the 5 per cent level of confidence. This significant correlation is explained in part by the control of intensity during the recording of the subject's responses for intelligibility testing. In testing the forty Spanish and Japanese subjects, sound pressure level was not controlled, thus, tending to cause the obtained intelligibility scores to be unreliable.

In the administration of the intelligibility test to the
thirty-six international students, sound pressure level was controlled, thus, increasing the reliability of the obtained scores and in turn enhancing the validity for the correlation between Tape I (nine-second passages) and intelligibility.

**The Reliability of Tape I (nine-second passages) for International Students.** In Chapter III, methods were described for establishing the reliability of Tape I (nine-second passages). The procedure for establishing reliability was based on the responses of native speakers of English. In order to test the reliability of Tape I (nine-second passages) in terms of non-native speakers, the responses for the twenty native speakers of Japanese and twenty speakers of Spanish were analyzed. Tape I (nine-second passages) is composed of sixteen passages. These passages had been ranked by criterion judges on a scale of one to eight. Each scale value was represented by two passages, thus, scale value 1 was represented by two passages, scale value 2 was represented by two passages, and so on for the eight scale values and sixteen passages.

As in the case for the native speakers of English, the mean responses for the forty non-native speakers of English were determined. The mean responses for each of the pairs of passages representing the eight scale values were correlated. The correlation was, \( r = .82 \). Corrected for length by the Spearman-Brown prophecy formula, the split-half
reliability coefficient was, $r = .90$. This indicated that this test is highly reliable for non-native speakers (.90) as well as for native speakers (.97).

On the basis of the evidence presented above, the hypothesis that judgments of controlled speech patterns are not reliable or valid estimates of intelligibility can be rejected insofar as concerns Tape I (nine-second passages).
CHAPTER V

CONCLUSIONS

This study was designed to test four null hypotheses concerning English as a foreign language. The hypotheses under test were:

1. There is no difference between native and non-native speakers of English (a) in the ability to predict the letters of English, (b) in the ability to comprehend English, (c) in the ability to differentiate between the sounds of English, (d) in intelligibility when speaking English, and (e) in the ability to judge controlled patterns of speech.

2. There is no difference between native speakers of Spanish and Japanese in the variables that are listed above.

3. There is no relationship among the variables that are listed above.

4. Judgments of controlled speech patterns are not reliable or valid estimates of intelligibility.

Tests of these hypotheses involved (a) selection and preparation of the materials to be presented to the subjects, (b) the selection of the experimental subjects, (c) administration of the materials to the subjects, (d) the application
of statistical techniques to assess the differences and relationships among the scores obtained for the various tests.

Selection and Preparation of the Materials for Presentation to the Subjects. Three of the measures selected for testing of the subjects had been designed for other investigations and were adopted for use in this study. These three tests were: (1) Lado's Test of Aural Comprehension. This test was developed at the English Language Institute, University of Michigan. Short statements are read to the students and they choose their answers from pictures and sentences. (2) Templin's Short Test of Sound Discrimination.\(^1\) In this test the subjects listen to pairs of nonsense syllables and indicate whether the members of the pairs are the same or different. (3) Form B of the Waco twenty-four word multiple-choice intelligibility test.\(^2\)

\(^1\) Mildred Templin, op. cit.

\(^2\) C. H. Haagen, op. cit.

The usual procedure for this test calls for the subjects to read lists of words to a panel of listeners. The listeners select the words that they think are spoken from multiple-choice items on a special answer sheet. In this investigation, the subject's responses were recorded as they read lists of words for the intelligibility test and these
responses were later played to five panels each composed of twenty native speakers of English. The instructions and stimulus materials for the first two tests mentioned above were recorded and played to the subjects at a constant intensity level.

Two tests were designed for the purpose of this investigation to assess speech proficiency indirectly. Fifty-four speakers composed of native and non-native speakers were recorded as they read four-word phrases equated on the basis of frequency of occurrence of individual words. The responses were edited so that nine-seconds of speech composed of at least five four-word phrases were retained for each speaker. The passages were rated on the basis of pronunciation excellence by a criterion group of 100 American college students (Group A). From the composite judgments of the judges, sixteen passages with the smallest semi-interquartile ranges (two each for eight scale values) were selected and arranged in a tape for presentation to the experimental subjects. This tape was designated as Tape I (nine-second passages). In the main study the judgments of the experimental subjects were compared with the scale values assigned by the criterion group. Total deviations from the values originally assigned by the criterion group were taken as test scores.
Two hundred and seventy four-word phrases selected from the responses of the fifty-four native and non-native speakers of English were also played to Group A. They were asked to indicate each word not spoken the way it should be spoken in English. Total judgments for each word were taken as an index of the pronunciations of each phrase. Phrases rated as relatively "good" in terms of pronunciation were paired with phrases rated as "poor" in pronunciation, first for native speakers of English, and then for non-native speakers of English. This resulted in a tape composed of thirty pairs of phrases. This tape was played to a group of graduate students in speech who were asked to judge which of the members of the pair were better pronounced. Their responses indicated a high degree of agreement with the values assigned to the members of the pairs. When this tape was played to a group of foreign students it was found that their scores were significantly lower than those for the American group. This tape was designated as Tape III (pairs).

A modification of Shannon's method\(^3\) was employed to

\(^3\) Shannon, loc. cit.

measure letter prediction ability. Answer sheets were coded for six phrases of fifteen letters each by means of a solution of potassium thio-cyanate and water. When the solution dried no clue to the presence of the solution
could be detected. To find the letters, the subjects used green ink containing ferric sulfate. When the correct letter was found the green ink turned brown. Each answer sheet contained fifteen columns composed of the letters of the orthographic alphabet plus a symbol for a space between words. Each subject marked the letters of the first column until the ink turned color. With the knowledge of the first letter of the phrase, he attempted to guess the letter in the second column, and so on, for the remaining letters in the phrase.

Selection of the Experimental Subjects. Twenty native speakers of Spanish, and twenty native speakers of English were selected as subjects for the first part of this study. Native speakers of Japanese and Spanish were chosen because (1) within each group the subjects have similar language backgrounds and encounter similar problems in understanding English. (2) Spanish is relatively similar to English, and (3) Japanese is quite unlike both Spanish and English.

Hypothesis I

There is no difference among native speakers of English, Spanish, and Japanese in comprehension of English, intelligibility, speech-sound discrimination, preferences for controlled speech patterns, and the ability to predict the letters of English.
In order to test this series of hypotheses, $t$ ratios for tests of significance among the means of scores obtained for the sixty subjects were computed. With the single exception of the scores for sound discrimination, the means for the English-speaking groups were significantly different at the 1 to 5 per cent level of confidence from the means obtained for the Japanese- and Spanish-speaking groups.

The means for letter prediction were based on the total trials for each letter for each subject counting no more than ten trials for each letter. The number of letters that was guessed correctly on the first trial was counted for each group. The $t$ ratios for the significance of differences of the means indicated significant differences between the native and non-native speakers of English. A $t$ ratio for the test of significant difference of the means for the number of letters predicted correctly after ten or more previous trials was significant at the 5 per cent level of confidence only for the English and Spanish groups. The difference was in favor of the Spanish group.

The hypothesis that there is no difference among native speakers of English, Spanish, and Japanese in comprehension of English, intelligibility, preferences for controlled speech patterns, and the ability to predict the letters of English was rejected with confidence insofar as the differences between native and non-native speakers of
English was concerned. No significant difference was found between the American and Japanese subjects and the Japanese and Spanish subjects at the 5 per cent level of confidence for the test of speech-sound discrimination. A significant difference was found for this test at the 1 per cent level of confidence for the English- and Spanish-speaking subjects.

**Hypothesis II**

**THERE IS NO DIFFERENCE BETWEEN NATIVE SPEAKERS OF SPANISH AND JAPANESE IN THE ABILITY TO PREDICT THE LETTERS OF ENGLISH, IN THE ABILITY TO COMPREHEND ENGLISH, IN THE ABILITY TO DIFFERENTIATE BETWEEN THE SOUNDS OF ENGLISH, IN INTELLIGIBILITY WHEN SPEAKING ENGLISH, AND IN THE ABILITY TO JUDGE CONTROLLED PATTERNS OF SPEECH.**

A significant difference between the means for the native speakers of Japanese and Spanish for the variables that are listed above was not found at the 5 per cent level of confidence. The hypothesis that there is no difference, therefore, was not rejected.

**Hypothesis III**

**THERE IS NO RELATIONSHIP AMONG SCORES OBTAINED FOR A TEST OF AURAL COMPREHENSION, A TEST OF SPEECH-SOUND DISCRIMINATION, TWO MEASURES DESIGNED TO TEST THE ABILITY TO JUDGE CONTROLLED SPEECH-PATTERNS, A TEST OF INTELLIGIBILITY, AND A TEST OF THE ABILITY TO PREDICT THE LETTERS OF ENGLISH.**
No significant product-moment correlations were found among the tests for the native speakers of English.

Significant correlations for the Spanish-speaking subjects were found between the test of aural comprehension and Tape I (nine-second passages), aural comprehension and Tape III (pairs) and aural comprehension and letter prediction ability. The product-moment correlations were respectively, .50, significant at the 5 per cent level of confidence, .59, significant at the 1 per cent level of confidence, and .49, significant at the 5 per cent level of confidence.

Significant product-moment correlations were found for the Japanese subjects between aural comprehension and letter prediction (.48) and sound discrimination and Tape I (nine-second passages) (.45). Both product-moment correlations were significant at the 5 per cent level of confidence.

When the scores for the six tests for the Japanese- and Spanish-speaking subjects were pooled, four significant correlations were found: aural comprehension and Tape I (nine-second passages) (.47), aural comprehension and letter prediction (.51) and Tape III (pairs) and Intelligibility (.36). The first two pairs of variables correlated at the 1 per cent level of confidence and the latter pair at the 5 per cent level of confidence.
Sound discrimination correlated with only one other variable, that is, Tape I (nine-second passages) at the 5 per cent level of confidence and only for the Japanese group.

The hypothesis that there is no relationship among scores for a test of aural comprehension, two measures designed to test the ability to judge controlled speech patterns, a test of intelligibility, and a test of the ability to predict the letters of English can be rejected with confidence as regards the non-native speakers. The hypothesis cannot be rejected that there is no relationship between the test of speech-sound discrimination and any other test, except in the one instance where the scores obtained for the test of speech-sound discrimination were correlated with scores for Tape I (nine-second passages) for the Japanese group.

**Hypothesis IV**

Judgments of controlled speech patterns are not reliable or valid estimates of intelligibility.

Tape I (nine-second passages) and Tape III (pairs) were designed to test pronunciation ability indirectly. Upon administration of Tape III (pairs) to the sixty subjects it was found that the range of scores for this measure was limited to ten errors. Because of this, this test was not tested further to determine its reliability or validity.
Tape I (nine-second passages) was administered to a second group of thirty-six international students none of whom were native speakers of English. These students were also administered the Waco twenty-four word write-down intelligibility test. The correlation between the scores obtained for the intelligibility test and Tape I (nine-second passages) was, \( r = .33 \), significant at the 5 per cent level of confidence.

The importance of this correlation as a coefficient of validity for the use of Tape I (nine-second passages) as an indirect test of pronunciation ability is enhanced by the previously established validity of this test. In Chapter II, the procedure for comparing Tape I (nine-second passages) with the judgments of the pronunciation excellence of the words in the passages was summarized. The resulting correlation was .84, significant at the 1 per cent level of confidence. When the scores for forty-two native speakers of English for Tape I (nine-second passages) were computed, a correlation of, \( r = .97 \) (split-half, corrected for length by the Spearman-Brown prophecy formula) was determined. This indicates high reliability for this test when administered to native speakers of English. When the mean responses for forty non-native speakers of Spanish and Japanese were analysed a product-moment coefficient of correlation of, \( r = .90 \) (split-half, corrected for length by the Spearman-Brown prophecy formula) was obtained. This indicates the high
reliability of this test for non-native speakers of English.

The validity of Tape I (nine-second passages) is based on the fact that the test is composed of a range of samples of English spoken by people who range from relatively good speakers of English to relatively poor speakers of English. Further, the values for the words in Tape I (nine-second passages) were determined independently and compared with ranks assigned to the sixteen passages in the test by the criterion group. A product-moment correlation of .84 was obtained, indicating that the passages in Tape I varied in pronunciation excellence. Coupled with the high reliability of this test, (r = .97, split-half corrected by the Spearman-Brown prophecy formula for native speakers of English and r = .90 for non-native speakers) the correlation of, r = .33 takes on added significance.

On the basis of the evidence, the hypothesis, insofar as Tape I (nine-second passages) is concerned, can be rejected that judgments of controlled speech patterns are not reliable or valid estimates of intelligibility.

Certain conclusions can be drawn from the results of this study. These conclusions are as follows:

1. Comprehension of English, as measured by Lado's test of aural comprehension is related to speech intelligibility.
2. As far as the native speakers of Spanish and Japanese selected for this study are concerned, intelligibility is not affected by language background.

3. Tape I (nine-second passages) is a valid and reliable indirect test of speech intelligibility.

4. Shannon's statement can be modified to read: "Anyone knowing a language possesses an enormous knowledge of the statistics of a language and this knowledge varies with his comprehension of the language."

These conclusions are discussed at greater length below.

1. COMPREHENSION OF ENGLISH, AS MEASURED BY LADO'S TEST OF AURAL COMPREHENSION, IS RELATED TO SPEECH INTELLIGIBILITY.

A test of aural comprehension of English and a test of intelligibility for speech were administered to twenty native speakers of Spanish and twenty native speakers of Japanese. When the scores obtained for the test of aural comprehension and those obtained for the test of speech intelligibility were compared for the pooled group of forty subjects the product-moment coefficient of correlation was .36, significant at the 5 per cent level of confidence. As intensity was not controlled when the Waco twenty-four word multiple-choice intelligibility test was used, a second group of thirty-six international students was tested. In this
instance, the Waco twenty-four word write-down intelligibility test was employed. The resultant correlation for intelligibility and aural comprehension was, $r = .42$, significant at the 1 per cent level of confidence. This leads to the conclusion that the more a person knows a language the more he will be understood by a native speaker. This generalization is true at least for the population under study in this investigation.

2. AS FAR AS THE NATIVE SPEAKERS OF SPANISH AND JAPANESE SELECTED FOR THIS STUDY ARE CONCERNED, INTELLIGIBILITY IS NOT AFFECTED BY LANGUAGE BACKGROUND.

No significant difference was found when the means obtained for the twenty native speakers of Japanese and twenty native speakers of Spanish for the test of aural comprehension and the test of the ability to predict the letters of English were compared. Thus, in terms of comprehension of English, the two groups were not essentially different. Also, the means obtained for the two groups of subjects for the test of intelligibility were not significantly different.

Because intensity was not controlled for the test of intelligibility in this part of the study, a legitimate question can be raised as to the reliability of the test. In spite of this, the possibility that relative intelligibility is the same for groups of international students from
different language backgrounds, raises questions relative to teaching methods. One of the important aspects of foreign-language learning is that the student must learn to speak so as to be understood by the native speaker of a language. As was shown above, the better one understands a language, the better one is understood by the native speaker of a language. Although training in phonetic accuracy aimed at speech-sound errors typical of speakers of various languages is important in improving intelligibility, it may be that this type of training is effective only when a concurrent effort is made to increase comprehension of the second language.

3. TAPE I (NINE-SECOND PASSAGES) IS A VALID AND RELIABLE INDIRECT TEST OF INTELLIGIBILITY.

A product-moment coefficient of correlation of .33, significant at the 5 per cent level of confidence, was obtained when the scores for thirty-six international students were compared for Tape I (nine-second passages) and the Waco twenty-four word write-down intelligibility test. The correlation indicated that the scores for the two tests tended to vary directly.

This statistically significant correlation is further enhanced when the evidence relative to reliability and validity of Tape I (nine-second passages) is examined.
This evidence is summarized in this chapter under Hypothesis IV.

4. SHANNON'S STATEMENT CAN BE MODIFIED TO READ: "ANYONE KNOWING A LANGUAGE POSSESSES AN ENORMOUS KNOWLEDGE OF THE STATISTICS OF A LANGUAGE AND THIS KNOWLEDGE VARIES WITH HIS COMPREHENSION OF THE LANGUAGE."

A correlation of, \( r = .51 \), significant at the 1 percent level of confidence was found when the total scores for the test of letter-prediction ability were compared with the scores obtained for aural comprehension for forty native speakers of Japanese and Spanish. On this basis it can be said that the ability to predict the letters of English tends to vary with comprehension of English.
BIBLIOGRAPHY
BIBLIOGRAPHY


Directory of International Students, Autumn Quarter, 1952. The Committee on International Students, 104 Administration Building, The Ohio State University, Columbus 10, Ohio.


*English Examination for Foreign Students*. Educational Testing Service, P. O. Box 592, Princeton, New Jersey.
Examination in Structure. English Language Institute, University of Michigan, Ann Arbor, Michigan.


Test of Aural Comprehension. English Language Institute, University of Michigan, Ann Arbor, Michigan.


Trubetzkoy, N. S., "Wie Soll das Lausystem einer kunstlichen internationalen Hilfssprache beschaffen sein?" *Travaux du Cercle linguistique de Prague*, 8:5-21, 1939.


APPENDIX A

ANSWER SHEET, ANSWER BOOKLET, AND EXAMINER'S MATERIAL FOR TEST OF AURAL COMPREHENSION
TEST OF AURAL COMPREHENSION
Form A

Instructions

ENGLISH. The examiner will read a sentence or a paragraph. You will select the one answer (A, B, or C) in the test which you consider correct and put a check (✓) beside the corresponding letter on the answer sheet. Do not mark the test. The examiner will read some examples.

ESPAÑOL. El examinador leerá una oración o un párrafo. Escoja la respuesta (A, B, o C) que juzgue correcta en la prueba y marque (✓) la letra correspondiente en la hoja especial que se le dará para tal efecto. No marque la prueba. El examinador leerá varios ejemplos.

PORTUGUÊS. O examinador deverá ler a sentença ou parágrafo. Deverá escolher a resposta (A, B, ou C) que pensa estar correta no test e marcar ao lado da letra correspondente na folha especial para respostas com um sinal (✓). Não marque no test. O examinador deverá ler alguns exemplos.

FRANÇAIS. L'examineur lira une phrase ou un paragraphe. Choisissez la réponse (A, B, ou C) que vous croyez être correcte dans le test et marquez (✓) la lettre correspondante sur la feuille destinée aux réponses, mais ne marquez rien sur le test lui-même. L'examineur donnera quelques exemples.
EXAMPLE: (A) in the living room.
   (B) at 6 o'clock.
   (C) if she studies.

1. (A) Mr. and Mrs. Smith have 4 pork chops for dinner.
   (B) Mr. and Mrs. Smith have 4 children.
   (C) Mr. and Mrs. Smith have invited 4 people to dinner.

2. (A) the home.
   (B) the school.
   (C) the father.

3. (A) Charles had 10 friends in the class.
   (B) Charles had an examination with 10 questions.
   (C) Charles had 10 errors on the examination.

4. (A) Mary likes the movies.
   (B) Mary accepted an invitation to the movies.
   (C) Mary does not like the movies.

5. (A) John had seen the accident.
   (B) John wanted to see the accident.
   (C) John did not see the accident.

6. (A) There are seats for approximately 50 people.
   (B) There are seats for more than 50 people.
   (C) There are seats for exactly 50 people.

7. (A) Mary told us the number of her house.
   (B) Mary told us the classes she attends.
   (C) Mary told us where she goes to church.

8. (A) There are many people walking on Main Street.
   (B) There are many red lights on Main Street.
   (C) There are many automobiles on Main Street.

9. (A) John paid a bill.
   (B) John obtained money.
   (C) John left his hat at the desk.

10. (A) A boy gave a book to a girl.
    (B) A girl gave a book to a boy.
    (C) A girl gave a book to a girl.

11. (A) Mary likes to swim.
    (B) Mary does not like to swim.
    (C) Mary cannot learn to swim.

12. (A) John does not know Detroit.
    (B) John is not going to Detroit.
    (C) John does not know if he is going to Detroit.

13. (A) John will go to the movies later.
    (B) John goes to the movies frequently.
    (C) John has finished his lessons.
14. (A) Mary asks something.
(B) Mary affirms something.
(C) Mary negates something.

15. (A) John did not know how to drive.
(B) John obtained a permit.
(C) John was careless in driving the car.

16. (A) The car went very fast.
(B) The car stopped.
(C) The car went over the hill.

17. (A) The price became smaller.
(B) The price became greater.
(C) The price remained the same.

18. (A) The decision was not easy.
(B) The decision was not difficult.
(C) The decision was not just.

19. (A) Mr. Smith has no money.
(B) Mr. Smith has a little money.
(C) Mr. Smith has much money.

20. (A) Mr. Smith paid the bill.
(B) Mr. Smith has no money.
(C) Mr. Smith did not pay the bill.

21. (A) This ticket does not permit me to stop anywhere.
(B) This ticket permits me to stop somewhere over the week-end.
(C) There is no refund on this ticket.

22. (A) John rarely goes to church.
(B) John frequently goes to church.
(C) John always goes to church.

23. (A) I am not happy.
(B) John sent me roses.
(C) John sent me a 5 pound box of candy.

24. (A) The boys are early.
(B) The boys are certainly going to be late.
(C) The boys do not have much time.

25. (A) The man was honest.
(B) The man was punctual.
(C) The man was temperate.

26. (A) There was an accident.
(B) I read about the development of airplanes.
(C) I read about a new airplane.

27. (A) It is time to go.
(B) We went at the right time.
(C) The weather is good.
28. (A) The weather is very cold.
   (B) We cannot endure the weather.
   (C) The weather is unusual.

29. (A) The principal actor was good.
   (B) The story in the movie was good.
   (C) There was an argument in the movie.

30. (A) Mary asked who sat there.
   (B) Mary asked me if I enjoyed sitting there.
   (C) Mary asked if there was any objection to her sitting there.

31. (A) I have forgotten how to sing and play.
   (B) I don't know if she sings or plays.
   (C) I always knew she sang and played the piano.

32. (A) The chairman invited suggestions.
   (B) The chairman never listened to any suggestions.
   (C) The chairman disregarded the wishes of the others.

33. (A) John was sincere.
   (B) John told the truth.
   (C) John did not tell the truth.

34. (A) Uncle Charles played pool every night.
   (B) Uncle Charles never played pool at night.
   (C) Uncle Charles didn't like to play pool.

35. (A) John went to the second hotel.
   (B) John went to a hotel room.
   (C) John went to a moderate hotel.

36. (A) I was disillusioned.
   (B) I was not disillusioned.
   (C) He was a good singer.

37. (A) We want to buy a new watch.
   (B) We like to see cargo ships pass.
   (C) We like to see freight trains pass.

38. (A) Mrs. Smith is healthy.
   (B) Mrs. Smith is good.
   (C) Mrs. Smith is rich.

39. (A) Mary selected the books carefully.
   (B) Mary selected the books without much thought or care.
   (C) Mary took the books from the floor.

40. (A) 2
   (B) 1
   (C) 3 years
TEST OF AURAL COMPREHENSION

Form A

Examiner's Material

(The examiner will read the following material to the students allowing enough time after each item for them to record their answers. The length of the pause should be between 12 and 15 seconds. The examiner should read in a clear natural style; no effort should be made to increase or decrease the difficulty of each item.)

Distribute the test booklets and the answer sheets to the students. Have them write their name and the date on the answer sheet only. Have them check the form of the test (A, B, or C) that is being given. Read the following instructions to them:

"I am going to read a sentence or a paragraph. You will select the one answer (A, B, or C) which you think is correct and put a check beside the corresponding letter on the answer sheet.

Example 1:

There is a circle, a triangle, and a square on your paper. Which is the triangle? (Pause)

The correct answer is "C". Check "C" on the answer sheet. Are there any questions?

Example 2:

The man is playing the violin. (Pause)

The correct answer is "B". (See that they are marking the proper place on the answer sheet).

Example 3:

The man is sitting on the chair.

The correct answer is "B".

Copyright, 1946
Charles C. Fries, Director and Robert Lado, Assistant Director
English Language Institute
University of Michigan
Ann Arbor, Michigan
Example 4:

He shut his suitcase by sitting on it.

The correct answer is "C".

Open your booklet to page 2. Do not interrupt. I will not repeat what I read.

PART I

1. The mother says, "John, please close the window."
2. John said, "Did you find the ball?" Charles said, "I'm still looking for it."
3. It's a beautiful day, and he is going fishing.
4. The man played a concert on the new piano at 7:30. When did he play?
5. It's his hat that fell in the water, not he.
6. He is going to eat with his sister.
7. A nice young man took off his hat in the elevator at 9:00 o'clock. Who took off his hat?
8. The dog is farther from the ball than the cat.
9. You're a lucky boy. She saw your picture and was very happy.
10. The boys had the bird tied with a string.
11. The boy is as tall as the girl.
12. Somebody turned the radio on too loud. Will you turn it down a little?
13. He wasn't able to maintain his balance.
14. The dog was bitten by the man.
15. He is giving her his hat and coat.
16. The fat man with the cigar is having his house painted.
17. The girl used the telephone.
18. The little boy likes milk chocolate.
19. She is sinking.
20. The horse's eyes were cold and piercing.

PART II

The only difference in Part II is that the answers are in writing instead of in pictures. "Mary eats dinner at 6 o'clock. She likes to study in the living-room after dinner. At what time does Mary eat dinner?" (Pause)

The correct answer is "B". Are there any questions?

(If a blackboard is available, the examiner may use additional examples by writing the following choices on the blackboard and reading the items below them.)

Write on blackboard:

Example 2: A. His mother wrote a letter.  
B. John wrote a letter.  
C. John and his mother wrote a letter.

Example 3: A. The train leaves at 6 o'clock.  
B. My friend lives in New York.  
C. The New York train leaves at 3 o'clock.

Read: "Example 2: John wrote a letter to his mother." (Pause)

"The correct answer is 'B'."

"Example 3: The train to New York leaves at three o'clock." (Pause)

"The correct answer is 'C'."

(Repeat the examples if necessary. Test items, however, may not be repeated.)

1. Mr. and Mrs. Smith have 4 guests for dinner.
2. The boy goes to school. The father goes to work. The mother stays at home. Who goes to work?
3. One day Charles had an examination. There were 100 questions, but Charles had only 10 mistakes.

4. John said, "Would you like to go the movie?" Mary said, "Yes."

5. John was asked if he had seen the accident. He said he hadn't.

6. There are seats in this room for about 50 people.

7. Mary told us her address.

8. The traffic is heavy on Main Street.

9. John cashed a check today.

10. A boy was walking down the street when he met a girl he knew. They stopped and talked for 10 minutes. She gave him a new book.

11. Mary is very fond of swimming.

12. John does not know whether he is going to Detroit or not.

13. When John has finished his lessons, he is going to the movies.

14. Mary said, "Is it convenient to go to New York by plane?" (falling intonation)

15. John got a license to drive the car.

16. John's car stalled on a hill.

17. The price of butter increased.

18. The decision of the judge was not fair.

19. Mr. Smith has a lot of money.

20. Mr. Smith neglected to pay the bill.

21. There is no stop-over on this ticket.

22. John seldom goes to church.

23. If John had sent me a 5 pound box of candy, I would not have been as happy as I am with these roses which he sent.
24. John says to Charles, "Hurry up, or we'll be late."
25. Mrs. Smith read a story about a man who never told a lie.
26. I read about an airplane crash.
27. It is time we went.
28. The weather is unseasonable.
29. The movie had a good plot.
30. Mary said, "Would you mind if I sat here?"
31. I've forgotten, if I ever knew, whether she sings or plays the piano.
32. The chairman overrode the requests of the other 2 committee members.
33. Had John been sincere he would have told me the truth.
34. There wasn't a night in the week Uncle Charles didn't go to play pool.
35. John went to a second-rate hotel and asked for a room.
36. If I had expected to hear a good singer, I would have been disappointed.
37. We like to watch the freighters go by.
38. Mrs. Smith is very well-to-do.
39. Mary picked 3 books at random.
40. John was in Mexico 3 years. It was his first and last visit to that country. How many times has John been to Mexico?
APPENDIX B

STIMULUS SYLLABLES AND ANSWER SHEET FOR
TEST OF SOUND DISCRIMINATION
### APPENDIX B: Stimulus Syllables for Templin's Short Test of Sound Discrimination

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>te-te</td>
<td>26.</td>
<td>ze-ze</td>
<td>51.</td>
</tr>
<tr>
<td>2.</td>
<td>hwe-we</td>
<td>27.</td>
<td>mai-nai</td>
<td>52.</td>
</tr>
<tr>
<td>3.</td>
<td>ne-me</td>
<td>28.</td>
<td>Ɂe-Ɂe</td>
<td>53.</td>
</tr>
<tr>
<td>4.</td>
<td>ge-de</td>
<td>29.</td>
<td>he-he</td>
<td>54.</td>
</tr>
<tr>
<td>5.</td>
<td>fi-vi</td>
<td>30.</td>
<td>dʒi-jı</td>
<td>55.</td>
</tr>
<tr>
<td>6.</td>
<td>he-pe</td>
<td>31.</td>
<td>pe-ke</td>
<td>56.</td>
</tr>
<tr>
<td>7.</td>
<td>se-ze</td>
<td>32.</td>
<td>tʃo-ʃo</td>
<td>57.</td>
</tr>
<tr>
<td>8.</td>
<td>Ɂe-Ɂe</td>
<td>33.</td>
<td>ki-ti</td>
<td>58.</td>
</tr>
<tr>
<td>9.</td>
<td>ʒe-dʒe</td>
<td>34.</td>
<td>eb-eb</td>
<td>59.</td>
</tr>
<tr>
<td>10.</td>
<td>vo-bo</td>
<td>35.</td>
<td>ehwe-ewe</td>
<td>60.</td>
</tr>
<tr>
<td>11.</td>
<td>ne-ne</td>
<td>36.</td>
<td>en-em</td>
<td>61.</td>
</tr>
<tr>
<td>12.</td>
<td>ʒe-tʃe</td>
<td>37.</td>
<td>ed-ed</td>
<td>62.</td>
</tr>
<tr>
<td>13.</td>
<td>ʃe-tʃe</td>
<td>38.</td>
<td>ehe-epe</td>
<td>63.</td>
</tr>
<tr>
<td>14.</td>
<td>im-ιy</td>
<td>39.</td>
<td>ov-ov</td>
<td>64.</td>
</tr>
<tr>
<td>15.</td>
<td>hwı-wı</td>
<td>40.</td>
<td>eʔ-eʔ</td>
<td>65.</td>
</tr>
<tr>
<td>16.</td>
<td>ge-ge</td>
<td>41.</td>
<td>eʃ-edʒ</td>
<td>66.</td>
</tr>
<tr>
<td>17.</td>
<td>dʒi-tʃi</td>
<td>42.</td>
<td>ov-ob</td>
<td>67.</td>
</tr>
<tr>
<td>18.</td>
<td>fai-fai</td>
<td>43.</td>
<td>ed-ed</td>
<td>68.</td>
</tr>
<tr>
<td>19.</td>
<td>ʒe-ve</td>
<td>44.</td>
<td>en-en</td>
<td>69.</td>
</tr>
<tr>
<td>20.</td>
<td>pe-pe</td>
<td>45.</td>
<td>ed-εt</td>
<td>70.</td>
</tr>
<tr>
<td>21.</td>
<td>fo-fo</td>
<td>46.</td>
<td>eʃ-εʃ</td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>vo-ʃo</td>
<td>47.</td>
<td>imi-imı</td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>so-so</td>
<td>48.</td>
<td>ihwı-iwı</td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td>ʃe-ʃe</td>
<td>49.</td>
<td>ɐɡ-ɐɡ</td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td>fi-ʃi</td>
<td>50.</td>
<td>ıa-ıa</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B: Directions for Subjects for Templin's Short Test of Sound Discrimination

You will hear pairs of syllables. If you think the syllables in the pair are the same, write S on your answer sheet. If you think the syllables in the pair are different, write D on your answer sheet. For example, you may hear (te-de). If you think both of the syllables are the same you will write S on your paper. If you think they are different, you will write D on your paper.
Appendix B: Answer Sheet for Templin’s Short Test of Sound Discrimination

<table>
<thead>
<tr>
<th>PAGE</th>
<th>FORM</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>DATE</th>
</tr>
</thead>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>31</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>34</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>37</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>40</td>
<td>41</td>
<td>42</td>
</tr>
<tr>
<td>43</td>
<td>44</td>
<td>45</td>
</tr>
<tr>
<td>46</td>
<td>47</td>
<td>48</td>
</tr>
<tr>
<td>49</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>52</td>
<td>53</td>
<td>54</td>
</tr>
<tr>
<td>55</td>
<td>56</td>
<td>57</td>
</tr>
<tr>
<td>58</td>
<td>59</td>
<td>60</td>
</tr>
<tr>
<td>61</td>
<td>62</td>
<td>63</td>
</tr>
<tr>
<td>64</td>
<td>65</td>
<td>66</td>
</tr>
<tr>
<td>67</td>
<td>68</td>
<td>69</td>
</tr>
<tr>
<td>70</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>73</td>
<td>74</td>
<td>75</td>
</tr>
</tbody>
</table>

161
APPENDIX C

ANSWER SHEET AND WORD LISTS FOR WACO MULTIPLE-CHOICE INTELLIGIBILITY TEST (FORM B)
WHEN YOU HEAR....

Number 1
MORTAR SHUT ASSIST
Number 2
BLIMP INJURE KNOB
Number 3
GLIDING BATTLE IGNITE

MARK YOUR PAPER LIKE THIS....
<table>
<thead>
<tr>
<th>Speaker</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker 1</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
<td>7th</td>
<td>8th</td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
<td>12th</td>
</tr>
<tr>
<td>Speaker 2</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
<td>7th</td>
<td>8th</td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
<td>12th</td>
</tr>
<tr>
<td>Speaker 3</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
<td>7th</td>
<td>8th</td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
<td>12th</td>
</tr>
<tr>
<td>Speaker 4</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
<td>7th</td>
<td>8th</td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
<td>12th</td>
</tr>
<tr>
<td>Speaker 5</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
<td>7th</td>
<td>8th</td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
<td>12th</td>
</tr>
<tr>
<td>Speaker 6</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
<td>7th</td>
<td>8th</td>
<td>9th</td>
<td>10th</td>
<td>11th</td>
<td>12th</td>
</tr>
</tbody>
</table>

Note: The table contains placeholder text and may not represent actual content.
APPENDIX C: Lists for Waco Multiple-choice Intelligibility Test (Form B)

Speaker 1

Number 1 piston firm banner
Number 2 eve attain scream
Number 3 rupture tour medal
Number 4 ark spotter gain
Number 5 cannon detract made
Number 6 lumber case pierce
Number 7 jail glimmer ward
Number 8 nature enact old

Speaker 2

Number 1 muzzle carve author
Number 2 scorch able cloth
Number 3 vision fumble groan
Number 4 cape lecture high
Number 5 possess blow single
Number 6 divide fiction maker
Number 7 leaf section rich
Number 8 traitor eastward join

Speaker 3

Number 1 traffic woolen swim
Number 2 can pulp eldest
Number 3 tank promote apt
Number 4 bad defend slight
Number 5 formal unpack license
Number 6 socket find misuse
Number 7 quarter kit rummage
Number 8 lock bond grind

Speaker 4

Number 1 tender beg swift
Number 2 wise nothing shark
Number 3 map full observe
Number 4 pace list contrast
Number 5 unheard just grow
Number 6 range hungry fade
Number 7 confine boast white
Number 8 naval race discard
Speaker 5

Number 1  year prone forest
Number 2  stable blast visit
Number 3  oval shock cost
Number 4  tack vital mar
Number 5  glimpse bomb untied
Number 6  dike salute mock
Number 7  fort dance pirate
Number 8  nest release cold

Speaker 6

Number 1  splash clock lag
Number 2  attack flavor gloom
Number 3  clap week this
Number 4  pause yet southeast
Number 5  drop arrest leather
Number 6  court fall navy
Number 7  bandit vast shatter
Number 8  word blame needy

Speaker 7

Number 1  tree pressure own
Number 2  cannon surf tonight
Number 3  ramp mold choose
Number 4  bench cancel arch
Number 5  senate got plunge
Number 6  harm flight drain
Number 7  fact throat it
Number 8  air gate outing

Speaker 8

Number 1  trench repay stiff
Number 2  rabbit exempt shield
Number 3  coach former preside
Number 4  blunder quaint worm
Number 5  purr dental bed
Number 6  keys good disturb
Number 7  aspire chute help
Number 8  germ await seaside
Speaker 9

Number 1  call handle vacant
Number 2  dust wallet leave
Number 3  fruit bail pants
Number 4  dish sitting jolly
Number 5  replace hunter bath
Number 6  journal fluid none
Number 7  plate cabin acute
Number 8  suffer clown grim

Speaker 10

Number 1  rock try ceiling
Number 2  drawer creep wool
Number 3  funnel nap lure
Number 4  shake hanger crop
Number 5  zone intend throw
Number 6  morning duck art
Number 7  inn kennel false
Number 8  belly shell bin

Speaker 11

Number 1  salt crash furnace
Number 2  moment ride broom
Number 3  clipper door tie
Number 4  sink matter half
Number 5  carbon end tone
Number 6  us passage front
Number 7  rate punch hotel
Number 8  dart bill at

Speaker 12

Number 1  raider view fear
Number 2  screen man beaver
Number 3  tense gun achieve
Number 4  captain salad drive
Number 5  instance brass helpful
Number 6  sweep drill meantime
Number 7  grasping file wave
Number 8  click wrench throb
APPENDIX D

ANSWER SHEET FOR LETTER PREDICTION TEST
APPENDIX E

DIRECTIONS FOR SUBJECTS AND ANSWER SHEETS FOR TAPE II
(FOUR-WORD PHRASES)
APPENDIX E: Directions for Subjects for Tape II (four-word phrases)

Now you will hear phrases followed by short pauses. For example:

pull ( )
the ( )
stick( )
back ( )

Put a check in the parenthesis for each word you think is not spoken the way it should be spoken in English. For example, if you think that pull and the are spoken well and stick and back are spoken poorly, put a mark in the parenthesis after stick and back.
### Appendix E: Answer Sheets for Tape II (four-word phrases)

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 1. if the engine fails | 10. back into the field | 19. bank into the field | 22. keep the plane from |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 2. hold back of neutral | 11. the plane will stall | 20. one thousand feet up | 23. cross wind | 26. turn the field |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 3. point; to land short | 12. to a full landing | 21. cross wind | 24. to a full landing | 27. a clear landing turn | 30. an airplane reaches |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 4. as in the approach | 13. you touch the ground | 22. starting to break it | 25. to circle the field | 28. can you always make a clear landing space | 31. in their proper space |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 5. so I'll take over | 14. by flying; down wind | 23. land into the field | 26. down to a field | 29. if the plane is landed | 32. the plane is landed |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 6. a clear landing space | 15. the turn is started | 24. to a full landing | 27. a clear landing turn | 30. an airplane reaches |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 7. are seated in it | 16. you touch the ground | 25. to circle the field | 28. can you always make a clear landing space | 31. in their proper space |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 8. point; to land short | 17. if there is a little | 26. down to a field | 29. if the plane is landed | 32. the plane is landed |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 9. the plane put | 18. he can always make a clear landing space | 27. a clear landing turn | 30. an airplane reaches |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**Abbreviations:**
- bank (bank): bank;
- field (field): field;
- plane (plane): plane;
- from (from): from;
- land (land): land;
- take off (take off): take off;
- stall (stall): stall;
- reach (reach): reach;
Appendix E: (cont.)

37. into the turn
38. before the plane lands
39. won't have to go
40. so that we'll have
41. we're in the air
42. the first landing shot
43. so that we'll have
44. never turn your back
45. a strong cross wind
46. late in the turn
47. and into the wind

48. wings are seated in it
49. speed is kept level
50. is flying across wind
51. don't have to go
52. off the back pro-nuro
53. one using room down
54. make a gliding turn
55. by fling down wind
56. starting to break it
57. the in started
58. and into the wind

59. are seated in it
60. cross wind takes off
61. you can always miss
62. apply to that shot
63. hold back of neutral
64. nuro that you miss
65. nuro that you miss
66. hold the glide longer
67. with the feeling that
68. land into the field
69. try to call
70. apply to that shot
71. if there is little
72. than a good one
73. try one small field
74. than a good one
75. hold the glide longer
76. with the feeling that
77. in the anyone!
78. to circle the hold
79. the plane will stall
80. try one small field
<table>
<thead>
<tr>
<th>61. the plane is put in the air</th>
<th>92. you're flying down wind</th>
<th>103. at the proper time</th>
<th>114. on the other hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>62. you're in the air</td>
<td>93. push the stick forward</td>
<td>104. in their proper space</td>
<td>115. we increase the rate</td>
</tr>
<tr>
<td>63. speed is kept constant</td>
<td>94. at the proper time</td>
<td>105. off the back pressure</td>
<td>116. feeling of the plane</td>
</tr>
<tr>
<td>64. the first landing shot</td>
<td>95. is flying across wind</td>
<td>106. before the plane lands</td>
<td>117. clear of the other plane</td>
</tr>
<tr>
<td>65. make the second one</td>
<td>96. the plane is landing</td>
<td>107. you're over the field</td>
<td>118. bank in the acre</td>
</tr>
<tr>
<td>66. soon as we clear</td>
<td>97. distance from the field</td>
<td>108. come then you're crossing</td>
<td>119. for the acre reason</td>
</tr>
<tr>
<td>67. wings are kept level</td>
<td>98. never turn your back:</td>
<td>109. try some turns now</td>
<td>120. break: the landing slide</td>
</tr>
<tr>
<td>68. a strong cross wind</td>
<td>99. you can follow no</td>
<td>110. should make your landings</td>
<td>121. you have a great</td>
</tr>
<tr>
<td>69. distance from the field</td>
<td>100. you're over the field</td>
<td>111. feeling of the plane</td>
<td>122. the next talk off</td>
</tr>
<tr>
<td>70. if we were coming</td>
<td>101. you can follow no</td>
<td>112. speed of the plane</td>
<td>123. you first cross it</td>
</tr>
<tr>
<td>71. we're flying down wind</td>
<td>102. push the stick forward</td>
<td>113. stick should be held</td>
<td>124. you have a great</td>
</tr>
</tbody>
</table>
Appendix E: (cont.)

125. we begin a turn
126. you have no power
127. bank is the same
128. mix the second one
129. clear of the field
130. to start to level
131. loadings are put on
132. break the landing glide
133. the field is roughly
134. if the engine fails
135. on the other hand

136. if there is little
137. gives you greater speed
138. try some circle shots
139. keep a constant bank
140. keep the landing point
141. ve circle the field
142. only then can we
143. try some circle shots
144. the degree of bank:
145. the stick back more
146. straight along the ground

147. to have the student
148. the degree of back
149. in a normal turn
150. back: pressure is hold
151. feel of the plane
152. head into the course
153. ve got out farther
154. the stick back more
155. crossing the wind line
156. as the other plane
157. ve level the wings
158. back: to level flight
159. to have the student
160. ve back: pressure
161. keep the plane from:
162. ve circle the field
163. the next take off
164. to start to level
165. you first cross it
166. crossing the wind line
167. crossing the wind line
168. ve increase the rate
169. ve level the wings
170. clear of the field
<table>
<thead>
<tr>
<th>169. you have no power</th>
<th>170. we correct for this</th>
<th>171. should make your landings</th>
<th>172. the plane is climbing</th>
<th>173. the plane is climbing</th>
<th>174. at the same time</th>
<th>175. up too much speed</th>
<th>176. landings are put on</th>
<th>177. stick should be held</th>
<th>178. we correct for this</th>
<th>179. we begin a turn</th>
</tr>
</thead>
<tbody>
<tr>
<td>180. the number of turns</td>
<td>181. you must be careful</td>
<td>182. the number of turns</td>
<td>183. at the same time</td>
<td>184. up too much speed</td>
<td>185. speed of the plane</td>
<td>186. the field is roughly</td>
<td>187. for the same reason</td>
<td>188. you must be careful</td>
<td>189. clear of other planes</td>
<td>190. in a normal turn</td>
</tr>
<tr>
<td>191. we level the wings</td>
<td>192. back pressure is held</td>
<td>193. cross the wind line</td>
<td>194. only then can we</td>
<td>195. if the other planes</td>
<td>196. we get out farther</td>
<td>197. give you good practice</td>
<td>198. make a good landing</td>
<td>199. we leave the traffic</td>
<td>200. we leave the traffic</td>
<td>201. for a normal glide</td>
</tr>
<tr>
<td>202. back to level flight</td>
<td>203. turn into the wind</td>
<td>204. keep the landing point</td>
<td>205. give you good practice</td>
<td>206. head into the course</td>
<td>207. start breaking the glide</td>
<td>208. make a good landing</td>
<td>209. start breaking the glide</td>
<td>210. for a normal glide</td>
<td>211. turn into the wind</td>
<td>212. keep a constant bank</td>
</tr>
</tbody>
</table>
Appendix B: (cont.)

213. use back pressure

214. straight along the ground

215. make a good landing

216. the degree of bank

217. never turn your back

218. the glider is broken

219. such a way that

220. see what you think

221. start to level out

222. one two three method

223. lower the right wing

224. such a way that

225. that you can land

226. you what I mean

227. quarter of a turn

228. it's a climbing turn

229. feel of the plane

230. will cause the plane to climb

231. road to fly alone

232. the degree of bank

233. you that I mean

234. closer to the field

235. even if you see

236. crosses the wind line

237. try to turn back

238. lower the right wing

239. that the student sees

240. it's a climbing turn

241. you have flying speed

242. try to turn back

243. one two three method

244. road to fly alone

245. plane starts to climb

246. road to fly alone

247. mean the heading is straight

248. lower the right wing

249. try to turn back

250. closer to the field

251. one thousand feet up

252. the glider is broken

253. in for the landing

254. in for the landing

255. to a poor surface

256. that is being done
Appendix E: (cont.)

257. that ( )
     the
     student
     noes
     
258. start
     to
     level
     put
     
259. one
     thousand
     foot
     up
     
260. even
     if
     you
     noe
     
261. can
     start
     our
     turn
     
262. by
     a
     diving
     turn
     
263. can
     start
     our
     turn
     
264. to
     a
     poor
     surface
     
265. that
     in
     being
     done
     
266. too
     that
     you
     think:
     
267. you
     have
     flying
     good
     
268. quarter
     of
     the
     turn
     
269. by
     a
     diving
     turn
     
270. that
     you
     can
     land
     

APPENDIX F

WORD LISTS AND ANSWER SHEET FOR WACO WRITE-DOWN INTELLIGIBILITY TEST
### APPENDIX F: Word Lists for Waco Write-down Intelligibility Test

<table>
<thead>
<tr>
<th>List 1</th>
<th>List 2</th>
<th>List 3</th>
<th>List 4</th>
<th>List 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>rupture</td>
<td>menace</td>
<td>zero</td>
<td>mute</td>
<td>outline</td>
</tr>
<tr>
<td>behold</td>
<td>pun</td>
<td>royal</td>
<td>city</td>
<td>uncle</td>
</tr>
<tr>
<td>transfer</td>
<td>half</td>
<td>converge</td>
<td>improve</td>
<td>improve</td>
</tr>
<tr>
<td>habit</td>
<td>salad</td>
<td>sixteen</td>
<td>shiftless</td>
<td>oblige</td>
</tr>
<tr>
<td>replace</td>
<td>navy</td>
<td>daily</td>
<td>quack</td>
<td>icebox</td>
</tr>
<tr>
<td>peanut</td>
<td>keyboard</td>
<td>propose</td>
<td>in</td>
<td>virtue</td>
</tr>
<tr>
<td>finance</td>
<td>owl</td>
<td>urgent</td>
<td>gallant</td>
<td>pressure</td>
</tr>
<tr>
<td>journal</td>
<td>dismay</td>
<td>entry</td>
<td>visit</td>
<td>divide</td>
</tr>
<tr>
<td>jar</td>
<td>made</td>
<td>jail</td>
<td>bishop</td>
<td>nature</td>
</tr>
<tr>
<td>garden</td>
<td>southeast</td>
<td>flicker</td>
<td>ampera</td>
<td>author</td>
</tr>
<tr>
<td>option</td>
<td>measure</td>
<td>candy</td>
<td>western</td>
<td>sweet</td>
</tr>
<tr>
<td>discard</td>
<td>darken</td>
<td>felt</td>
<td>income</td>
<td>volley</td>
</tr>
<tr>
<td>horror</td>
<td>headache</td>
<td>year</td>
<td>perfume</td>
<td>opal</td>
</tr>
<tr>
<td>function</td>
<td>journey</td>
<td>binding</td>
<td>ransom</td>
<td>recent</td>
</tr>
<tr>
<td>gang</td>
<td>rich</td>
<td>knuckle</td>
<td>faithfiul</td>
<td>hall</td>
</tr>
<tr>
<td>naval</td>
<td>lady</td>
<td>goat</td>
<td>wicked</td>
<td>irk</td>
</tr>
<tr>
<td>startle</td>
<td>world</td>
<td>fable</td>
<td>quiver</td>
<td>vacant</td>
</tr>
<tr>
<td>maiden</td>
<td>preserve</td>
<td>disgust</td>
<td>tonight</td>
<td>distract</td>
</tr>
<tr>
<td>fence</td>
<td>woolen</td>
<td>beef</td>
<td>air</td>
<td>minute</td>
</tr>
<tr>
<td>us</td>
<td>beware</td>
<td>envy</td>
<td>double</td>
<td>deafen</td>
</tr>
<tr>
<td>acorn</td>
<td>beaver</td>
<td>junction</td>
<td>held</td>
<td>laundry</td>
</tr>
<tr>
<td>tiger</td>
<td>galore</td>
<td>late</td>
<td>odd</td>
<td>pebble</td>
</tr>
<tr>
<td>concern</td>
<td>end</td>
<td>entrust</td>
<td>common</td>
<td>blow</td>
</tr>
<tr>
<td>find</td>
<td>pulp</td>
<td>pad</td>
<td>rubber</td>
<td>fact</td>
</tr>
<tr>
<td>List 6</td>
<td>List 7</td>
<td>List 8</td>
<td>List 9</td>
<td>List 10</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>nasal</td>
<td>arrest</td>
<td>hip</td>
<td>forbid</td>
<td>ocean</td>
</tr>
<tr>
<td>verse</td>
<td>hardly</td>
<td>got</td>
<td>rabbit</td>
<td>prow</td>
</tr>
<tr>
<td>suspend</td>
<td>thirteen</td>
<td>lash</td>
<td>perfect</td>
<td>mar</td>
</tr>
<tr>
<td>heel</td>
<td>excite</td>
<td>merry</td>
<td>raider</td>
<td>hit</td>
</tr>
<tr>
<td>arrow</td>
<td>became</td>
<td>lanky</td>
<td>bird</td>
<td>decent</td>
</tr>
<tr>
<td>bill</td>
<td>across</td>
<td>impulse</td>
<td>join</td>
<td>fasten</td>
</tr>
<tr>
<td>beg</td>
<td>speed</td>
<td>pause</td>
<td>earmark</td>
<td>salt</td>
</tr>
<tr>
<td>goal</td>
<td>arch</td>
<td>rabble</td>
<td>helpful</td>
<td>motion</td>
</tr>
<tr>
<td>partner</td>
<td>neutral</td>
<td>misuse</td>
<td>young</td>
<td>himself</td>
</tr>
<tr>
<td>farflung</td>
<td>hungry</td>
<td>unit</td>
<td>purge</td>
<td>each</td>
</tr>
<tr>
<td>grenade</td>
<td>traitor</td>
<td>ablaze</td>
<td>relief</td>
<td>proud</td>
</tr>
<tr>
<td>robust</td>
<td>project</td>
<td>manner</td>
<td>vacuum</td>
<td>wine</td>
</tr>
<tr>
<td>ice</td>
<td>epic</td>
<td>design</td>
<td>plate</td>
<td>carbon</td>
</tr>
<tr>
<td>living</td>
<td>rub</td>
<td>inland</td>
<td>chapter</td>
<td>icebound</td>
</tr>
<tr>
<td>gamble</td>
<td>sitting</td>
<td>up</td>
<td>often</td>
<td>enclose</td>
</tr>
<tr>
<td>former</td>
<td>medal</td>
<td>acute</td>
<td>rusty</td>
<td>yet</td>
</tr>
<tr>
<td>help</td>
<td>usage</td>
<td>oppose</td>
<td>dot</td>
<td>target</td>
</tr>
<tr>
<td>forego</td>
<td>race</td>
<td>reform</td>
<td>thimble</td>
<td>engage</td>
</tr>
<tr>
<td>prevail</td>
<td>eve</td>
<td>earthquake</td>
<td>vermin</td>
<td>old</td>
</tr>
<tr>
<td>weather</td>
<td>pace</td>
<td>upright</td>
<td>leaf</td>
<td>yellow</td>
</tr>
<tr>
<td>enquire</td>
<td>wallet</td>
<td>fourteen</td>
<td>address</td>
<td>fair</td>
</tr>
<tr>
<td>guilty</td>
<td>odor</td>
<td>jawbone</td>
<td>crowd</td>
<td>absent</td>
</tr>
<tr>
<td>quaint</td>
<td>ego</td>
<td>content</td>
<td>wife</td>
<td>nonsense</td>
</tr>
<tr>
<td>art</td>
<td>infect</td>
<td>firm</td>
<td>hive</td>
<td>outlet</td>
</tr>
</tbody>
</table>
APPENDIX F: (continued)

<table>
<thead>
<tr>
<th>List 11</th>
<th>List 12</th>
<th>List 13</th>
<th>List 14</th>
<th>List 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>hotel</td>
<td>airport</td>
<td>lame</td>
<td>surprise</td>
<td>massive</td>
</tr>
<tr>
<td>forget</td>
<td>only</td>
<td>quart</td>
<td>abrupt</td>
<td>thud</td>
</tr>
<tr>
<td>hello</td>
<td>fifteen</td>
<td>errand</td>
<td>hope</td>
<td>itch</td>
</tr>
<tr>
<td>autumn</td>
<td>enroll</td>
<td>jobber</td>
<td>fishing</td>
<td>dental</td>
</tr>
<tr>
<td>rigid</td>
<td>conclude</td>
<td>profess</td>
<td>secure</td>
<td>ceiling</td>
</tr>
<tr>
<td>elm</td>
<td>invite</td>
<td>tropic</td>
<td>lumber</td>
<td>ox</td>
</tr>
<tr>
<td>lemon</td>
<td>ramp</td>
<td>zone</td>
<td>obstruct</td>
<td>gracious</td>
</tr>
<tr>
<td>useless</td>
<td>endorse</td>
<td>money</td>
<td>man</td>
<td>mad</td>
</tr>
<tr>
<td>issue</td>
<td>swear</td>
<td>tomb</td>
<td>ride</td>
<td>precede</td>
</tr>
<tr>
<td>yonder</td>
<td>calm</td>
<td>feast</td>
<td>quarter</td>
<td>law</td>
</tr>
<tr>
<td>copy</td>
<td>belief</td>
<td>enjoy</td>
<td>zigzag</td>
<td>halfback</td>
</tr>
<tr>
<td>beneath</td>
<td>profit</td>
<td>garage</td>
<td>quit</td>
<td>bright</td>
</tr>
<tr>
<td>leave</td>
<td>narrow</td>
<td>profane</td>
<td>hate</td>
<td>achieve</td>
</tr>
<tr>
<td>behalf</td>
<td>scatter</td>
<td>willow</td>
<td>blunder</td>
<td>damp</td>
</tr>
<tr>
<td>possess</td>
<td>wheat</td>
<td>yard</td>
<td>talent</td>
<td>resume</td>
</tr>
<tr>
<td>vain</td>
<td>throb</td>
<td>shipping</td>
<td>rain</td>
<td>you</td>
</tr>
<tr>
<td>seaside</td>
<td>mood</td>
<td>glide</td>
<td>sustain</td>
<td>below</td>
</tr>
<tr>
<td>thorn</td>
<td>keystone</td>
<td>quarrel</td>
<td>license</td>
<td>iceman</td>
</tr>
<tr>
<td>eldest</td>
<td>juice</td>
<td>younger</td>
<td>threat</td>
<td>ready</td>
</tr>
<tr>
<td>noon</td>
<td>moss</td>
<td>row</td>
<td>tassel</td>
<td>out</td>
</tr>
<tr>
<td>joyful</td>
<td>elsewhere</td>
<td>unload</td>
<td>born</td>
<td>wood</td>
</tr>
<tr>
<td>vibrate</td>
<td>joke</td>
<td>hazard</td>
<td>shoot</td>
<td>good</td>
</tr>
<tr>
<td>ask</td>
<td>nation</td>
<td>deduct</td>
<td>daisy</td>
<td>lifelike</td>
</tr>
<tr>
<td>case</td>
<td>keep</td>
<td>bat</td>
<td>oral</td>
<td>tender</td>
</tr>
</tbody>
</table>
### APPENDIX F: (continued)

<table>
<thead>
<tr>
<th>List 16</th>
<th>List 17</th>
<th>List 18</th>
<th>List 19</th>
<th>List 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>key</td>
<td>obey</td>
<td>screech</td>
<td>worship</td>
<td>knitting</td>
</tr>
<tr>
<td>zoom</td>
<td>without</td>
<td>dance</td>
<td>stampede</td>
<td>native</td>
</tr>
<tr>
<td>escape</td>
<td>final</td>
<td>this</td>
<td>silver</td>
<td>herald</td>
</tr>
<tr>
<td>meaning</td>
<td>acquaint</td>
<td>maker</td>
<td>trapeze</td>
<td>pain</td>
</tr>
<tr>
<td>offend</td>
<td>ugly</td>
<td>ark</td>
<td>call</td>
<td>kernel</td>
</tr>
<tr>
<td>faint</td>
<td>lamp</td>
<td>dainty</td>
<td>uphill</td>
<td>swoop</td>
</tr>
<tr>
<td>fashion</td>
<td>hangar</td>
<td>crop</td>
<td>urge</td>
<td>district</td>
</tr>
<tr>
<td>trifle</td>
<td>wallet</td>
<td>nothing</td>
<td>gather</td>
<td>rummage</td>
</tr>
<tr>
<td>parlor</td>
<td>banquet</td>
<td>raid</td>
<td>lava</td>
<td>grasping</td>
</tr>
<tr>
<td>timid</td>
<td>heading</td>
<td>induce</td>
<td>enough</td>
<td>acre</td>
</tr>
<tr>
<td>unclean</td>
<td>shut</td>
<td>beyond</td>
<td>kiss</td>
<td>culture</td>
</tr>
<tr>
<td>brutal</td>
<td>captain</td>
<td>legion</td>
<td>temper</td>
<td>traffic</td>
</tr>
<tr>
<td>thousand</td>
<td>cabin</td>
<td>persist</td>
<td>valve</td>
<td>recline</td>
</tr>
<tr>
<td>room</td>
<td>harvest</td>
<td>darling</td>
<td>germ</td>
<td>anger</td>
</tr>
<tr>
<td>killer</td>
<td>doe</td>
<td>treasure</td>
<td>pardon</td>
<td>oval</td>
</tr>
<tr>
<td>wade</td>
<td>manage</td>
<td>withhold</td>
<td>storage</td>
<td>rumor</td>
</tr>
<tr>
<td>valley</td>
<td>glimmer</td>
<td>jumbo</td>
<td>at</td>
<td>abroad</td>
</tr>
<tr>
<td>cannon</td>
<td>heat</td>
<td>wise</td>
<td>keel</td>
<td>bombard</td>
</tr>
<tr>
<td>transcribe</td>
<td>drab</td>
<td>reverse</td>
<td>fault</td>
<td>detail</td>
</tr>
<tr>
<td>open</td>
<td>zinc</td>
<td>direct</td>
<td>morn</td>
<td>walking</td>
</tr>
<tr>
<td>question</td>
<td>wool</td>
<td>figure</td>
<td>observe</td>
<td>hurdle</td>
</tr>
<tr>
<td>jockey</td>
<td>banner</td>
<td>occur</td>
<td>junior</td>
<td>index</td>
</tr>
<tr>
<td>accent</td>
<td>campaign</td>
<td>handcuff</td>
<td>molest</td>
<td>abound</td>
</tr>
<tr>
<td>oath</td>
<td>robe</td>
<td>vote</td>
<td>nap</td>
<td>boast</td>
</tr>
</tbody>
</table>
### List 21

- broken
- week
- foreclose
- extreme
- chance
- upon
- driven
- able
- golf
- ignite
- untold
- slaughter
- gearshift
- victor
- class
- describe
- foremost
- afraid
- rage
- census
- north
- union
- distance
- salute

### List 22

- vast
- shudder
- eighty
- crook
- justice
- mankind
- napkin
- practice
- complete
- soldier
- panic
- tackle
- kodak
- nature
- kingdom
- wall
- lament
- tack
- import
- canteen
- pistol
- needy
- orphan
- quake
Appendix F: Answer Sheet for Waco Write-down Intelligibility Test
APPENDIX G

SCORES FOR SIXTY NATIVE SPEAKERS OF ENGLISH, SPANISH, AND JAPANESE FOR SIX TESTS
APPENDIX G: Scores for Sixty Native Speakers of English, Spanish, and Japanese

<table>
<thead>
<tr>
<th></th>
<th>Aural Comprehension</th>
<th>Sound Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Spanish</td>
<td>Japanese</td>
</tr>
<tr>
<td>60</td>
<td>49</td>
<td>56</td>
</tr>
<tr>
<td>60</td>
<td>40</td>
<td>52</td>
</tr>
<tr>
<td>54</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>60</td>
<td>50</td>
<td>49</td>
</tr>
<tr>
<td>60</td>
<td>49</td>
<td>57</td>
</tr>
<tr>
<td>60</td>
<td>46</td>
<td>52</td>
</tr>
<tr>
<td>57</td>
<td>60</td>
<td>46</td>
</tr>
<tr>
<td>60</td>
<td>52</td>
<td>56</td>
</tr>
<tr>
<td>59</td>
<td>47</td>
<td>50</td>
</tr>
<tr>
<td>59</td>
<td>41</td>
<td>37</td>
</tr>
<tr>
<td>57</td>
<td>34</td>
<td>57</td>
</tr>
<tr>
<td>59</td>
<td>26</td>
<td>45</td>
</tr>
<tr>
<td>59</td>
<td>49</td>
<td>54</td>
</tr>
<tr>
<td>58</td>
<td>40</td>
<td>57</td>
</tr>
<tr>
<td>58</td>
<td>58</td>
<td>44</td>
</tr>
<tr>
<td>60</td>
<td>57</td>
<td>38</td>
</tr>
<tr>
<td>56</td>
<td>45</td>
<td>47</td>
</tr>
<tr>
<td>55</td>
<td>51</td>
<td>53</td>
</tr>
<tr>
<td>59</td>
<td>51</td>
<td>47</td>
</tr>
<tr>
<td>60</td>
<td>48</td>
<td>44</td>
</tr>
</tbody>
</table>
APPENDIX G: (continued) Scores for Sixty Native Speakers of English, Spanish, and Japanese

<table>
<thead>
<tr>
<th>Intelligibility</th>
<th>Letter Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Spanish</td>
</tr>
<tr>
<td>60.6</td>
<td>54.7</td>
</tr>
<tr>
<td>69.3</td>
<td>64.4</td>
</tr>
<tr>
<td>78.0</td>
<td>38.1</td>
</tr>
<tr>
<td>69.7</td>
<td>59.5</td>
</tr>
<tr>
<td>74.6</td>
<td>55.7</td>
</tr>
<tr>
<td>81.7</td>
<td>38.1</td>
</tr>
<tr>
<td>62.2</td>
<td>55.3</td>
</tr>
<tr>
<td>88.8</td>
<td>64.5</td>
</tr>
<tr>
<td>79.2</td>
<td>40.9</td>
</tr>
<tr>
<td>61.9</td>
<td>55.3</td>
</tr>
<tr>
<td>69.3</td>
<td>52.2</td>
</tr>
<tr>
<td>64.3</td>
<td>30.2</td>
</tr>
<tr>
<td>72.4</td>
<td>78.4</td>
</tr>
<tr>
<td>60.7</td>
<td>69.5</td>
</tr>
<tr>
<td>62.4</td>
<td>60.1</td>
</tr>
<tr>
<td>50.1</td>
<td>59.3</td>
</tr>
<tr>
<td>35.8</td>
<td>39.5</td>
</tr>
<tr>
<td>57.4</td>
<td>55.1</td>
</tr>
<tr>
<td>70.5</td>
<td>60.5</td>
</tr>
<tr>
<td>62.8</td>
<td>48.9</td>
</tr>
</tbody>
</table>
### APPENDIX G: (concluded) Scores for Sixty Native Speakers of English, Spanish, and Japanese

<table>
<thead>
<tr>
<th>Tape I (nine-second passages)</th>
<th>Tape III (pairs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Spanish</td>
</tr>
<tr>
<td>18</td>
<td>33</td>
</tr>
<tr>
<td>26</td>
<td>37</td>
</tr>
<tr>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>18</td>
<td>41</td>
</tr>
<tr>
<td>12</td>
<td>39</td>
</tr>
<tr>
<td>12</td>
<td>34</td>
</tr>
<tr>
<td>18</td>
<td>33</td>
</tr>
<tr>
<td>18</td>
<td>42</td>
</tr>
<tr>
<td>29</td>
<td>40</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>
APPENDIX H

SCORES FOR THIRTY-SIX NON-NATIVE SPEAKERS OF ENGLISH FOR THREE TESTS
APPENDIX H: Scores for Thirty-six Non-native Speakers of English for Three Tests

<table>
<thead>
<tr>
<th>Intelligibility</th>
<th>Aural Comprehension</th>
<th>Tape I (nine-second passages)</th>
</tr>
</thead>
<tbody>
<tr>
<td>54.5</td>
<td>56</td>
<td>22</td>
</tr>
<tr>
<td>66.1</td>
<td>50</td>
<td>28</td>
</tr>
<tr>
<td>80.3</td>
<td>52</td>
<td>41</td>
</tr>
<tr>
<td>84.1</td>
<td>57</td>
<td>24</td>
</tr>
<tr>
<td>84.4</td>
<td>52</td>
<td>25</td>
</tr>
<tr>
<td>40.8</td>
<td>58</td>
<td>18</td>
</tr>
<tr>
<td>54.1</td>
<td>57</td>
<td>22</td>
</tr>
<tr>
<td>84.9</td>
<td>60</td>
<td>12</td>
</tr>
<tr>
<td>81.5</td>
<td>57</td>
<td>26</td>
</tr>
<tr>
<td>64.5</td>
<td>48</td>
<td>25</td>
</tr>
<tr>
<td>36.6</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>63.6</td>
<td>44</td>
<td>34</td>
</tr>
<tr>
<td>56.2</td>
<td>49</td>
<td>33</td>
</tr>
<tr>
<td>79.4</td>
<td>44</td>
<td>34</td>
</tr>
<tr>
<td>72.3</td>
<td>52</td>
<td>21</td>
</tr>
<tr>
<td>49.5</td>
<td>53</td>
<td>36</td>
</tr>
<tr>
<td>59.5</td>
<td>54</td>
<td>24</td>
</tr>
<tr>
<td>59.5</td>
<td>52</td>
<td>42</td>
</tr>
<tr>
<td>80.0</td>
<td>47</td>
<td>35</td>
</tr>
<tr>
<td>87.8</td>
<td>56</td>
<td>24</td>
</tr>
<tr>
<td>74.5</td>
<td>52</td>
<td>25</td>
</tr>
<tr>
<td>68.6</td>
<td>53</td>
<td>37</td>
</tr>
<tr>
<td>75.7</td>
<td>56</td>
<td>22</td>
</tr>
<tr>
<td>59.5</td>
<td>57</td>
<td>24</td>
</tr>
<tr>
<td>73.6</td>
<td>45</td>
<td>25</td>
</tr>
<tr>
<td>63.2</td>
<td>57</td>
<td>33</td>
</tr>
<tr>
<td>71.1</td>
<td>53</td>
<td>26</td>
</tr>
<tr>
<td>54.5</td>
<td>52</td>
<td>28</td>
</tr>
<tr>
<td>56.2</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>38.3</td>
<td>34</td>
<td>43</td>
</tr>
<tr>
<td>60.3</td>
<td>43</td>
<td>27</td>
</tr>
<tr>
<td>54.9</td>
<td>46</td>
<td>41</td>
</tr>
<tr>
<td>63.2</td>
<td>45</td>
<td>19</td>
</tr>
<tr>
<td>80.7</td>
<td>52</td>
<td>26</td>
</tr>
<tr>
<td>81.9</td>
<td>59</td>
<td>20</td>
</tr>
<tr>
<td>50.7</td>
<td>56</td>
<td>22</td>
</tr>
</tbody>
</table>
I, Robert Leonard Mulder, was born in Grand Haven, Michigan, July 25, 1922. I received my secondary school education in the public schools of the city of Grand Haven, Michigan. My undergraduate training was obtained at Western Michigan College of Education, Kalamazoo, Michigan, from which I received the degree Bachelor of Arts in 1948. I received the degree Master of Arts from The Ohio State University in 1949. While in residence at The Ohio State University, I acted in the capacity of Assistant in the Department of Speech. I held this position while completing the requirements for the degree Doctor of Philosophy.