DOCTORAL DISSERTATION SERIES

TITLE AN EXPERIMENTAL STUDY OF CERTAIN RELATIONSHIPS BETWEEN INTELLIGIBILITY SCORES AND CLINICAL DATA OF PERSONS WITH DEFECTIVE ARTICULATION

AUTHOR VICTOR PAUL GARWOOD

UNIVERSITY OF MICHIGAN

DATE 1951

DEGREE Ph. D. PUBLICATION NO. 3748



UNIVERSITY MICROFILMS
ANN ARBOR · MICHIGAN

AN EXPERIMENTAL STUDY OF CERTAIN RELATIONSHIPS BETWEEN INTELLIGIBILITY SCORES AND CLINICAL DATA OF PERSONS WITH DEFECTIVE ARTICULATION

by
Victor P. Garwood

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigam
1951

Committee in charge:

Professor H. Harlan Bloomer, Chairman
Assistant Professor W. H. Beaven
Professor G.E. Densmore
Assistant Professor D.R. Miller
Assistant Professor D.E. Morley

ACKNOWLEDGEMENTS

A study such as this represents the work and efforts of many people. Without the cooperation of the following people and organizations, the study could never have been completed. They are presented in the order of their appearance during the planning and execution of the experiment.

Dr. Ira Hirsh, Dr. E.D. Schubert, Dr. Chas. M. Lightfoot, Dr. Milton Dickens and the staff of the Department of Speech, University of Southern California; Prof. William Sener and the staff of Station KUSC, Hancock Foundation, University of Southern California; Prof. Herbert Farmer and the staff of the Department of Audio Visual Education, University of Southern California; Prof. L.E. Travis and the staff of the Speech and Hearing Clinic, Department of Speech, University of Southern California; Miss Esther Herbert, Coordinator of Speech Correction, Secondary Level, and her staff in the public school system of the City of Los Angeles; and, of course, heartfelt thanks to the sixty-four subjects and the forty members of the listening crews.

V.P.G.

TABLE OF CONTENTS

Page

LIST OF TABLES. v

LIST OF ILLUSTRATIONS vi

Chapter

1. Introduction 1

1.1 Significance of the Problem
1.2 Definition of Terms
1.3 Review of the Literature
1.4 Statement of the Problem

2. Subjects, Materials, Procedures 16

2.01 The Overview
2.02 The Experimental Design
2.03 Selection of Speakers
2.04 Rating Scale
2.05 Test Materials
2.06 The Communication Link
2.07 Preparation of Speakers for Recording
2.08 Recording the Speaker
2.09 Preparation of the Recordings
2.10 Preparation of the Studio for Listening Crews.
2.11 Selection of Listeners
2.12 Presentation of Recordings to the Listeners
2.13 Scoring the Listener Data Sheets

3. Presentation and Analysis of Data 36

3.1 Relationships between CA Scores and Ratings
3.2 Relationships between the EI Scores and Ratings.
3.3 Relationships between the Diagnostic Test and the Experimental Test
3.4 Differences between the Speaker Categories
3.5 Sex Variance
3.6 Word List Variance
3.7 Row and Column Variance
3.8 Summary of the Interpretation of the Data

4. A Discussion of the Effects of Phonetic Error Upon the Perception of Speech 51

4.1 Overview

TABLE OF CONTENTS - Continued

Page

Chapter

4.2 Selection of a Population Sample
4.3 Preparation of material for Discussion
4.4 Discussion of EI and ESE Scores
4.5 Discussion of Error Differences and Similarities between Categories III and IV
4.6 Interpretation of the Sample Data

5. Summary, Conclusions and Implications . . . 61

5.1. Summary
5.2 Conclusions
5.3 Summary of Findings in Chapter 4
5.4 Implications

6. APPENDIX 70

7. BIBLIOGRAPHY 104

LIST OF TABLES

Table		Page
1.	The experimental Design	21
2.	EI Score and Rating Correlations for the Four Experimental Speech Categories . . .	37
3.	Correlations Between the Two Testing Techniques in Both Test Situations For Category III	39
4.	Correlations Between the Two Testing Techniques in Both Test Situations For Category IV.	40
5.	Analysis of Variance for the EI Scores. .	46
6.	Analysis of Variance for the EI Ratings .	46
7.	Averaged EI Scores and EI Ratings Converted to Intelligibility Scores and Intelligibility Scores in Rank Order	47
8.	Averaged Phonetic Content of All PBF Lists and Averaged ESE and EI Scores--All Speech Categories	53
9.	Phonetic Content of Form C-2 and Averaged CAE Scores for Categories III and IV. . .	55
10.	Phonetic Content of Form C-2 and the Four PBF Lists	71
11.	CA Scores and Ratings in The Diagnostic Articulation Test	73
12.	EI Scores and Ratings by Sessions. . . .	74
13.	Tabulation of Phonetic Errors Noted by Expert Examiners in Categories III and IV . . .	82
14.	Sums of Errors and Ratings by Sessions . .	84
15.	Sounds Missed by Listeners in the Sixteen Speaker Sample	88

LIST OF ILLUSTRATIONS

Figure		Page
1.	Plan of Studio "C"	27

CHAPTER 1

INTRODUCTION

1.1 Significance of the Problem

The first essential factor of oral communication is the ability of the speaker to make himself understood by his listener. For a successful reception of a message unit, the listener must use a set of phonemes that corresponds to the set of phonemes chosen by the speaker in that particular situation. Any aberrancy on the speaker's part increases the number of decisions to be made by the listener as to message content, reducing the efficiency of the entire communication process. It would, therefore, seem that any study which might throw light upon common speaker aberrancies might well be considered significant.

Speakers deviate from "normal" in regard to speech factors such as rate, rhythm, pitch, or such subjective voice aspects as "pleasantness," forcefulness," and the like. Any or all such deviations may impair the speaker-listener relationship. However, it has been generally assumed that articulation deviations occupy a "key" position since accurate reception of the intended sounds is the first phase of the communication process upon which all other phases depend. Therefore, an experimental study of

listeners' reactions to articulatory deviant speakers would seem to be significant.

Finally, it might be noted that many previous studies in the general area of obstacles to effective communication have produced important findings, and at the same time have indicated certain research "gaps." Such studies attest to the significance of further research in the area. Before reviewing the literature of previous research, several terms should be defined.

1.2 Definition of Terms

1.21 Defective Articulation was defined in this study as any deviation in articulation noted by an expert examiner, ie., phonemic omissions, distortions, and/or substitutions during attempted utterance of speech in the diagnostic test situation.

1.22 Pronunication was defined as "choice, arrangement, and accenting of sounds by a normally functioning articulatory mechanism."[1]

1.23 Intelligibility was defined in this study as the degree to which the words as spoken by a speaker were recognized (properly identified) by the listener as being the same as the written stimulus words in the experimental test situation.

(Early in the present study it was noted that a variety of related terms was found in the appropriate research

---

[1] H. Blattner, "an Experimental Study of the Testing of Pronunciation," Speech Monographs, XV (1948), p. 181.

literature. There were, for example, such terms as: "Articulation Test, "Intelligibility," etc. The usage of such terms by different researchers has often been overlapping or contradictory. Therefore, for the purposes of the present study, the following principal terms were arbitrarily chosen and operationally defined.)

1.24 Clinical Articulation Scores, or "CA Scores," were defined as the number of errors of articulation noted by the expert examiners during the diagnostic articulation test.

1.25 Clinical Articulation Ratings, or "CA Ratings," were defined as the expert examiners' subjective evaluation of the degree of severity of the articulation defect on a seven point rating scale.

1.26 Experimental Intelligibility Scores, or "EI Scores," were defined as the number of words written down by the naive listeners which were not the same as the written stimulus words uttered by the speaker in the experimental situation. For example, if the test word was "tap," and the speaker uttered the word correctly, all written responses other than "tap" such as "top," map," and "mat," were considered as errors and contributed along with other such errors to make up a total score. Thus, if the listener misidentified ten such words, his experimental score would be ten. From this score, an "Articulation Score" (or for purposes of this experiment, an Intelligibility score) based upon the correct responses could be computed, if necessary.

1.27 Experimental Intelligibility Ratings, or "EI Ratings," were defined as the listeners' subjective evaluation of the speaker's total utterances on a seven point rating scale, in terms of the listeners' "understanding" the speaker. This type of listener judgment has been often called an "Intelligibility Rating" by other researchers.

1.3 Review of the Literature

In general, problems of articulation and intelligibility may be conveniently classified into three categories, representing the three main situations in which speech and hearing training are currently practiced; the classroom, the clinic, and the laboratory. Professional persons most directly interested in the foregoing situations are respectively: the general speech teacher, the speech therapist, and the phonetician and audiologist.

1.31 The Laboratory

The audiologist and the audio-engineer have been interested primarily in the distortional effects produced by electronic communication equipment on the human ear, and generally speaking, the experimental procedure used normal speakers under various noise conditions.

Fletcher and Steinberg,[1] and French and Steinberg[2]

---

[1]H. Fletcher and J.C. Steinberg, "Articulation Testing Methods," Bell System Technical Journal, VIII (1929), 806-854.

[2]N.R. French and J.C. Steinberg, "Factors Governing the Intelligibility of Speech Sounds," Journal of the Acoustical Society of America, XIX (1947), 90-119.

of the Bell Telephone Laboratories devised a method whereby the intelligibility of speech could be estimated by means of proper physical measures of the communication and the noise conditions alone. Unfortunately for the speech pathologist, they used the term "Articulation Testing" instead of "Intelligibility Testing." They conceived of the telephone receiver as "articulating" to the listener. By using a series of words as test material they ran off a series of tests to find out if the telephone equipment was adequate for the ordinary purposes of telephonic communication.

During World War II this concept was expanded to envelop high fidelity sound transmission devices, especially under conditions of loud noise. Considerable experimentation under the auspices of the Office of Scientific Research and Development in collaboration with The Bell Laboratories, the Harvard Psycho-Acoustic Laboratory, and the armed forces resulted in a tremendous flow of information about the intelligibility of speech under fatiguing conditions. Eagan[1] has summarized the articulation test procedures, the history of testing, and test conditions generally employed in this type of research.

These studies were all geared to explore the variables encountered under wartime conditions and to discover how voice communication could be made more effective. The

---

[1]J.P. Eagan, "Articulation Testing Methods," Laryngoscope, LVIII (September, 1948), 955-991.

program has been well summarized by Black[1] and Steer and Hadley.[2] For example, the following were only a few of the questions the researchers sought to answer:

1) How should radio and intercom equipment in an airplane be handled or worn for greatest efficiency?[3]

2) How should one speak over the equipment in order to be as intelligible as possible?[4]

3) How may air crew members be trained efficiently in speaking and listening in order to attain maximum intelligibility?[5]

4) How may the habitual use of standard messages be taught?[6]

---

[1]J.W. Black, "The Origin and Nature of the Studies," (Studies in Speech Intelligibility), Speech Monographs, XIII (1946), 1-4.

[2]M. Steer and J.M. Hadley, "The Speech Intelligibility Program in Naval Aviation," Quarterly Journal of Speech, XXXII (April, 1946), 217-228

[3]J.F. Curtis, "Intelligibility Related to Microphone Position," Speech Monographs, XIII (1946), 8-13.

[4]F.L. Whan, "Training in Listening and in Voice and Diction for the Airplane Pilot," Quarterly Journal of Speech, XXX (October, 1944), 262-265.

[5]H.M. Mason, "The Effect of Very Loud Speech Signal Upon Intelligibility," Speech Monographs, XIII (1946), 19-24.

[6]H.M. Moser, "Intelligibility Related to Loudness," Speech Monographs, XIII (1946), 47-50.

5) How do pitch[1] and loudness[2] relate to intelligibility?

6) How do certain phonetic factors relate to the accuracy of speech in airplane noise?[3]

7) How does articulation relate to intelligibility?[4]

The last question, relating to articulation of speakers under noisy conditions, utilized normal and superior speakers as subjects. Draegert[5] concluded that training in good articulation patterns rather than training which attempted to improve the pronunciation of difficult sounds only, was likely to improve speech intelligibility.

Shortly after the war, the materials and procedures were released for publication and were utilized chiefly to estimate hearing thresholds and discrimination levels of the hard-of-hearing person. Typically, the procedure consisted of attenuating nonsense syllables, words or sentences, through a high fidelity sound system to the person being tested. The subjects used for these experiments consisted

---

[1] I.P. Brackett, "Intelligibility Related to Pitch," Speech Monographs, XIII, (1946), 24-32.

[2] P. Moore, "Intelligibility Related to Pitch," Speech Monographs, XIII, (1946), 13-19

[3] W.E. Stevens, "Study of Phonetic Factors in Relation to Accuracy of Transmission of Words in Airplane Noise," Speech Monographs, XIII (1946), 32-41.

[4] G.L. Draegert, "Intelligibility Related to Articulation," Speech Monographs, XIII (1946), 50-54.

[5] Ibid.

of a control group of normal listeners and an experimental group of hard-of-hearing patients. These tests, like those used by the armed forces, were more concerned with testing the listener, rather than testing the speaker. Haskins,[1] Quick[2] Hudgins et. al.,[3] and Falconer and Davis[4] have used modifications of the articulation test technique for this purpose.

Measurements of the intelligibility of defective speakers have been confined largely to the speech of the deaf. Hudgins and Numbers,[5] investigated the intelligibility of deaf pupils at the Clark and Airey schools. Using selected sentences as test materials, they had teachers judge speakers on a rating scale and compared those ratings with the tabulations of errors noted by two phonetic analyzers, they derived an "error index" from the number of sounds malarticulated and the frequency of occurrence in the speech

---

[1] H. Haskins, "A Phonetically Balanced Test of Speech Discrimination for Children," An Unpublished Master's Thesis, Northwestern University, (August, 1949), pp. 84 ff.

[2] M. Quick, "The Development of a Test of Speech Perception to Measure the Acoustic Achievement of Young Deaf Children," Unpublished Master's Thesis, Smith College, (June, 1949), pp. 51 ff.

[3] C.V. Hudgins, J.E. Hawkins, J.E. Karlin, S.S. Stevens, "The Development of Recorded Auditory Tests for Measuring Hearing Loss for Speech," Laryngoscope, LVII (January, 1947), 57-89.

[4] G.A. Falconer and H. Davis, "The Intelligibility of Connected Discourse as a Test for the Threshold for Speech," Laryngoscope, LVII (September, 1947), 581-595.

[5] C.V. Hudgins and F.C. Numbers, "An Investigation of the Intelligibility of the Speech of the Deaf," Genetic Psychology Monographs, XXV (1942), 289-392.

test material. They concluded that intelligibility was largely dependent upon the frequency of errors of articulation and errors of rhythm. Hudgins[1] later used the PBF (Phonetically Balanced, Familiar) lists and sentences on another group of deaf pupils and compared articulation scores of words and sentences. Using trained listeners, he found that a sentence score of 90% to 100% corresponded to a word score of approximately 50%, and also found that he could predict sentence scores from word scores with an error of approximately 10%.

Morley[2] compared articulation scores with sound spectrograms using the Bell Sound Spectrograph and concluded that the articulation function was important in the intelligibility of the speech of selected deaf speakers.

1.32 The Speech Clinic

Possibly, the speech correctionist or clinician has considered the aspect of intelligibility somewhat more intensively than the classroom teacher, simply because defective speakers were further down the speech continuum than the normal speaker--both qualitatively and quantitatively. In addition to possessing more specific articulatory aberrations, the speech defective person might exhibit certain behavior patterns which were more demonstrable clinically than

---

[1]C.V. Hudgins, "A Method of Appraising the Speech of the Deaf," Volta Revue, (December, 1949), 597-601; 638.

[2]D.E. Morley, "An Analysis by Means of the Sound Spectrograph of Intelligibility Variations of Consonant Sounds Spoken by Deaf Persons," (Ann Arbor: University Microfilms No. 1354, 1949).

the behavior patterns of the normal and "near-normal" speaker. Perhaps Van Riper's definition of a speech defective person would serve to define this particular category of speaker.

> "Speech is defective when it deviates so far from the speech of other people that it calls attention to itself, interferes with communication, or causes its possessor to be maladjusted."[1]

There seemed to be no specific studies extant in the literature of the intelligibility of persons with functional articulatory disorders or the so-called "Dyslalic" speakers. Lightfoot investigated the intelligibility of persons with "coryzal" speech, or speech affected by a transitory medical condition--the common cold, and concluded that "little or no relationship existed between the amount of influence catching a cold might have on a man's normal intelligibility and the distinctiveness of his coryzal speech."[2]

Other studies apparently disregarded the intelligibility factor in favor of studies to measure the degree of defective articulation. Curry et. al.,[3] used a series of recorded samples of speech of eight defective speakers and had listeners judge the records in terms of mild to severe,

---

[1] C. Van Riper, Speech Correction Principles and Methods, (2d ed., New York: Prentice-Hall, 1947) p. 15.

[2] C. M. Lightfoot, "Some Effects of the Common Cold on Speech," Archives of Otolaryngology, LI (April, 1950), 500-513.

[3] R. Curry, L. Kennedy, L. Wagner, and W. Wilke, "A Phonographic Scale for the Measurement of Defective Articulation," JSHD, VIII (1943), 123-126.

using the paired comparison technique. Certain scale steps were computed on the basis of judgments of the recorded samples. However, their problem was not one of assessing intelligibility per se, but one of measuring defective articulation. Wilke[1] also studied the "diction" of persons with a "New Yorkese" dialect and found that his "diction scale" indicated accurately a certain group of speech teachers' average opinions of the importance of certain deviations from preferred articulations.

Wood,[2] Henderson,[3] Bangs,[4] and Williams[5] dealt with the analysis of defective articulation more as relating to the quantity and quality of articulation errors than to speaker intelligibility. Similar studies were reported by Spriestersbach and Curtis[6] of work done more recently at the

---

[1]W.H. Wilke, "Development and Application of a Scale for Measuring Diction," Quarterly Journal of Speech, XXIV (1938), 268-281.

[2]K.S. Wood, "Parental Maladjustment and Functional Articulatory Defects in Children," Journal of Speech Disorders, XI (1946), 255-275.

[3]F.M. Henderson, "Accuracy in Testing the Articulation of Speech Sounds, "Journal of Educational Research, XXXI (1938), p. 348.

[4]J.L. Bangs, "A Clinical Analysis of the Articulatory Defects of the Feebleminded," Journal of Speech Disorders, VII (1942), 343-356.

[5]H.M. Williams, "A Qualitative Analysis of the Erroneous Speech Sound Substitutions of Pre-school Children," University of Iowa Studies in Child Welfare, XIII (1937), 19-32.

[6]D.C. Spriestersbach and J.F. Curtis, "Misarticulation and Discrimination of Speech Sounds," Quarterly Journal of Speech, XXXVII (December, 1951), 487-491.

State University of Iowa.

1.33 The Classroom

By contrast with the areas reviewed, comparatively little research of direct interest to the general classroom teacher has been reported. Teachers of speech have been concerned with intelligibility to the extent that, although not always mentioned as such, techniques were devised to aid the speaker to speak "more clearly," use the lips more," speak more distinctly," etc. Students were generally screened early in the academic year, and should the teacher spot some general speech inadequacy, those students were usually referred to a speech correctionist or clinician for further evaluation. If the inadequacy appeared to be chiefly articulatory in nature, the clinician decided whether the errors were specific and/or limited and possibly "pathological," or whether articulation in general was "inaccurate." If the clinical examination revealed no etiology of pathology and the inaccuracies were rather mild and/or non-specific, the clinician might suggest that the speech teacher work toward more general clarity in pronunciation or "enunciation." This category of speaker has generally been described as "slovenly," indistinct," mumbling," mouthing," lazy-lipped," or some like term connoting malarticulation. Apparently this type of speaker occupied a position along the general articulatory continuum close to the normal end, yet differentiated from the normal speaker--and at the same time not considered defective.

Latham,[1] in a discussion of faulty speech habits, referred to a statement in the "Introduction to High School Lessons" in the Teachers College Record; September, 1910, by Romiett Stevens:

> "This difficulty is further augmented by the fact of the almost universal custom of high school pupils to speak indistinctly, to pronounce words carelessly, or to trail off their sentences into nothingness."

Fay and Middleton,[2] compared "pleasantness of voice" in both the live voice and the electrically amplified voice situation on a seven point rating scale. They found that voices were less "pleasant" over a P.A. system, but also found that listeners' ratings of voice over a P.A. system were, in general, quite reliable.

It appeared, therefore, that the domain of the speech teacher lay in the areas of "normal" and, arbitrarily, "slovenly" speech as far as articulation was concerned. The judgments made were generally subjective in nature and usually reflected the experience and insight of the teacher.

Thus, these and other studies approached the subject of intelligibility either from the clinical standpoint of estimating hearing thresholds of the hard-of-hearing speaker; the intelligibility of the deaf person; or the intelligibility of the normal or superior speaker under simulated

---

[1] A.J. Latham, "Audibility and Distinctness in the Speech of High School Students," Quarterly Journal of Speech, XVII, (February, 1931), p. 2.

[2] P.J. Fay and W.C. Middleton, "Rating a Speaker's Natural Voice When Heard Over a Public Address System," Quarterly Journal of Speech, XXVII (February, 1941), p. 124.

wartime conditions. Estimates of normal, defective, and "slovenly" speakers under peacetime conditions have been highly subjective opinions of the expert examiner or the teacher of speech.

This study attempts to assess the intelligibility of speakers commonly instructed in speech classes, in the clinic, or in speech correction classes. It utilized techniques relating to both situations--subjective evaluation and diagnostic articulation testing. It uses the intelligibility testing procedure commonly used in testing the listener, which was ultimately derived from testing telephone circuits. In a sense, therefore, this study bridges the fields of the teaching of speech, speech pathology, and psycho-acoustics.

1.4 The Statement of the Problem

Generally speaking, this study attempted to answer this question: What relationships exist between certain experimental listener evaluations of intelligibility and certain clinical evaluations by expert examiners of persons with defective articulation?

Specifically, the problem may be stated in the form of the following five questions:

1) What relationships exist between the CA Scores and the CA Ratings in the diagnostic test situation?

2) What relationships exist between the EI Scores and the EI Ratings derived in the experimental test situation?

3) What relationships exist between the measurements made in the diagnostic test situation and the experimental test situation?

4) What relationships are expressed to differentiate the four chosen categories of speakers?

5) Does the sex of the listener in the experimental test situation produce differences in scores and ratings?

CHAPTER 2

SUBJECTS, MATERIALS, PROCEDURES

2.01 The Overview

In general, experimental procedures were as follows. Sixty-four selected male subjects, representing four different articulatory categories of speakers, were screened for errors of articulation. These speakers were recorded on magnetic tape, using Hudgins PBF word lists[1] as speech materials. The recordings were spliced in an order suggested by a Graeco-latin square design and then played back under controlled conditions to listening crews, so that each listening crew, composed of ten listeners, heard only sixteen speakers. Listeners were assigned two tasks: first, to write down the words that they heard the speaker utter; and secondly, to assign a numerical intelligibility score. Raw scores were tabulated and prepared for analysis. The foregoing procedures will next be described in greater detail.

2.02 The Experimental Design

2.021 The Diagnostic Test Situation

Two expert examiners, employed in the field of speech correction and/or pathology, were used to analyze and tabulate errors of articulation and to judge, independently, the severity of articulation defects. This procedure served to screen out any subjects selected who were not members of

---

[1]Hudgins, "A Method of Appraising the Speech of the Deaf," op. cit., p. 638.

the population of defective speakers. They also performed the task of screening the non-clinical population in order to screen out possible "defective" speakers. Each speaker was required to read aloud the same selection of reading material, during which time the examiners noted any errors of articulation. This procedure was followed by a subjective evaluation of the severity of the defect. This procedure will be more fully explained in Section 2.03

2.22 The Experimental Test Situation

In order to control as many possible sources of variation which might affect the speakers, lists, listeners, etc., in the experiment, test conditions were set up as follows:

1) Equal numbers of speakers were placed in each speech category to be tested.

2) Their word lists and randomized replications were equally distributed throughout the experiment so that each list would have an equal chance of being uttered by each category of speakers in the session.

3) All of the speakers uttered a standard phrase before the test word for monitoring purposes so that the phrase was a constant factor for all speakers, and at the same time regarded as a sample of connected discourse.

4) Every listening crew heard the same number of speakers and speaker categories heard by other listening crews.

5) Each listener in the listening crew had an

equal chance to hear each speaker.

6) The environment in the speaker-listener situation remained relatively constant so that all listening crews would have an equal chance to hear every speaker.

The Graeco-latin Square design[1] seemed adequate to fulfill the requirements just enumerated. In this type of design, treatments are grouped into replicates in three different ways with the consequence that the effects of the sources of variation are equalized for all treatments. Essentially the speakers were arranged in four squares. Each square represented a listener session and contained four categories of speakers and four randomizations of the test materials. From these replicated squares, two complete sets of data were derived; one set was composed of data relating to the experimental intelligibility scores, and the other to the experimental intelligibility ratings.

It was hypothesized that although listening crews would hear each list and each category four times, the arrangements of the speaker categories and the randomization of the test materials would offset any "learning" during the session. Hudgins concluded that

> "....familiarity with the word order of any list was avoided by employing different scramblings of each list--the structure of the lists is such that most of the words in lists have from one to six rhyme words, some of which appear in each of the other lists. Hence, familiarity with the

---

[1] W.G. Cochran and G.M. Cox, Experimental Designs, (New York: John Wiley, 1950), p. 117.

> words would be of little advantage since the correct word must be identified out of a group of similar or rhyming words....Since no pupil reads a list in an order identical to that read by any other pupil, the listeners never memorize the word order."[1]

Four different listening crews were selected to audit the speakers. The experimenter was not interested in "expert" listeners, but rather a sample of "untrained" people, i.e., those who would not be familiar with the technique nor the theory of intelligibility testing. They will be referred to in this study as "naive, "untrained listeners," or "listeners." These listeners were chosen from undergraduate speech courses because the experimenter assumed that they were sufficiently motivated to attend carefully to the task set for them.

Listening crews were split into equal male and female groups for each session to find out whether measurements of the variables were subject to sex error. Guilford[2] stated that in rating scale techniques men were inclined to be more lenient in their ratings than were women when judging certain human traits, and that there was a tendency to overrate members of the same sex as compared with members of the opposite sex. Ten members for each crew were considered as adequately reliable for fairly stable results. Eagan[3] reported that six to eight trained listeners would yield

---

[1] Hudgins, op. cit., p. 599

[2] J.P. Guilford, Psychometric Methods, (New York: McGraw Hill, 1936), p. 276.

[3] J.P. Eagan, "Transmission and Reception of Sounds Under Combat Conditions," Summary Technical Report of Division 17, NDRC, (OSRD III, Washington, D.C: 1946)Vol III, P.73.

fairly accurate results. Haagen[4] used eight to twelve members, who were both speakers and listeners. He stated that the predicted reliability for different numbers of listeners from eight to twenty varied only ten points and showed that in actual measurement the correlations tended to become higher.

Table 1 shows the complete experimental design by session, speakers, and test Materials.

TABLE 1

THE EXPERIMENTAL DESIGN

CODE: Speaker category in Roman numerals
Test materials in Capital letters
Randomization of test materials in Arabic numerals

SESSION 1.				SESSION 2.			
IA1	IIC2	IIID3	IVB4	IVD2	IIIC2	IIB3	IA4
IIB1	ID2	IVC2	IIIA4	IIIB1	IVA1	ID3	IIC3
IIIC1	IVA2	IB3	IID4	IIA1	IB2	IVC1	IIID4
IVD1	IIIB2	IIA2	IC4	IC1	IID2	IIIA3	IVB3
SESSION 3.				**SESSION 4.**			
IIIB3	IC2	IIA4	IVD4	IIC4	IIIB4	IVA4	ID4
IIC1	IVB2	IIID2	IA3	IA2	IVD3	IIIC3	IIB2
ID1	IIIA2	IVC3	IIB4	IIID1	IIA2	IB1	IVC4
IVA3	IID1	IB4	IIIC4	IVB1	IC3	IID3	IIIA1

[4]C.H. Haagen, "Intelligibility Measurement," Speech Monographs, XIII (1946), p. 6.

2.03. Selection of Speakers

The design provided for four different categories of speakers with varying degrees of articulatory proficiency. These categories might yield significantly different intelligibility scores. The writer wanted a homogenous group of subjects if possible. Since differences of pitch and voice quality might conceivably affect the selection only one sex was chosen. As the majority of clinic cases were males, it was decided to capitalize on this chance factor and restrict the samples to male speakers. It was also fortunate that most of the referrals from the International Center were also males. All four categories had no other clinically demonstrable speech disorders. Although the four groups represented only two populations, clinical and non-clinical, to the expert examiner. The four categories were classified as follows:

Category I. Sixteen male speakers whose speech had been classified as "normal" or "superior" by university speech teachers.

Category II. Sixteen male speakers whose speech had been classified as "slovenly" by high school, or university speech teachers.

Category III. Sixteen male speakers whose native language was not General American or English. (Foreign Accent group).

Category IV. Sixteen male speakers whose speech had

been classified as "Dyslalic" by expert clinical examiners.

Categories I and II were termed non-clinical, i.e., those speakers could not be considered for treatment as pathological cases. These speakers were selected by university and high school teachers of speech purely on a subjective basis, but subject to the following criteria:

1) Normal speakers were defined as those speakers whose articulation of sounds during the production of a test paragraph were free of any omissions, distortions, and/or substitutions, slurring, regional dialect, etc.

2) Slovenly speakers were defined as those speakers, who, by a process of exclusion from the above category, demonstrated that they were not normal speakers, had no traces of a foreign accent, and were not, nor have ever been cases for clinical consideration.

Following the selection, subjects in both categories read aloud a paragraph of reading material to a clinical examiner. (See Form C-3 in the Appendix). Speakers with demonstrable sound errors were excluded from these categories.

Subjects in category III were selected by the coordinator of the International Center, University of Southern California, from the restricted population of East Indian students in attendance at the University. This group, therefore, presented a very select sample of members whose native language was relatively homogenous and to whom the English language was a "second" or "acquired" language. They were taught the English language in India by British

speaking native Indian teachers. These speakers came from either the Calcutta or the Bombay area.

Subjects in category IV were selected from the case load of the Speech and Hearing Clinic, University of Southern California, and from speech correction classes in the Hollywood, Los Angeles, and Manual Arts Senior High Schools of Los Angeles. Eleven of these individuals were lispers, one had a (w) substitution for (r) and (l), and four had combinations of (s), (r), and (l) substitutions.

Although categories III and IV were the usual members of clinical populations, the former, strictly speaking, were members by a geographical accident! Categories I, II, and IV, were native-born citizens of the United States.

Subjects in categories III and IV were given a phonetic analysis by two speech clinicians consisting of a series of test words and sentences. These "expert examiners" were selected by the Clinic faculty staff. After the analysis of each individual, the examiners compared their tabulations and, when necessary, had them repeat certain test sentences until a perfect agreement of noted errors was reached. Then each clinician made an independent judgment of the speaker's defect severity, based on the test sentences, oral reading of the previously mentioned paragraph, and spontaneous speech. This judgment was made on a seven point rating scale designated as the "Severity of Articulation Defect," (Clinical Articulation Rating). The instruction sheet for the examiners (C-1) and the data sheet,

(C-2), are reproduced in the appendix.

All the subjects of the experiment were given a pure tone, audiometric sweep-check examination before being selected. None of the subjects had any demonstrable hearing loss when checked by approved audiometric methods.

2.04 Rating Scale

A seven point rating scale was used in the Diagnostic Test situation and in the Experimental Test situation. The former scale attempts to assess the continuum of defect severity, while the latter was concerned with the intelligibility continuum. The rationale for this scale over the perhaps more popular five point scale was taken from Symonds,[1] who concluded that the degree of reliability one wished in the final ratings should decide the matter. He stated that, for rating human traits, seven was the optimal number, as more than seven classes yielded an increase in reliability hardly great enough to pay for the trouble. Less than seven led to an appreciable loss in reliability.

For the purposes of this experiment, it was only possible to rate the speakers in categories III and IV in the Diagnostic Test situation as I and II had no demonstrable articulatory disorders.

2.05 Test Materials

The four PBF (Phonetically Balanced, Familiar) mono-

---

[1] P.M. Symonds, "On the Loss of Reliability in Ratings Due to the Coarseness of the Scale," J. Exper. Psychol., VII (1924), p. 461.

syllabic word lists, selected by Hudgins from the twenty PB lists used in the Harvard studies, were used. These four lists of fifty words each were selected by him as being identical in construction with respect to phonetic content with the original PB lists and differed only in that they were composed of familiar words, i.e., words found in elementary school readers.[1] Due to the varying degrees of schooling and the varied cultural backgrounds of the subjects, it was felt that such basic word material would give more stable results than the original lists. Each list was replicated in a different random order so that the four basic lists were expanded to sixteen lists. The nomenclature was changed somewhat to conform to the purposes of the experiment. List 1. became "List A" with its four randomizations, 1,2,3, and 4. The other three sets of lists were labelled B,C, and D, and treated similarly. This procedure made it possible for each subject in each category to use a different variation of the lists.

The phonetic content of the four PBF lists and the Diagnostic Articulation test is presented on Table 9 in the Appendix. The four word lists are also reproduced in the Appendix.

2.06 <u>The Communication Link</u>

High fidelity sound equipment was used throughout the study and all recordings and playbacks were handled in

---

[1]Personal letter from C.V. Hudgins, Director of Research, The Clarke School For the Deaf, May 2, 1951.

the same sound treated suite of rooms in Broadcasting Studio "C" in the Alan Hancock Foundation Building at the University of Southern California. This studio is an acoustically treated room 21 ft. 6 in. wide by 27 ft. long by 12 ft. 6 in. high with an insulated control room adjacent and visible to the studio. A Western Electric 639a Cardioid microphone, switched to the dynamic pattern setting, was placed in front of a small table and chair. The microphone was on the end of a moveable boom, which made it possible to move the microphone to any desired position. The microphone was jacked into the input stage of a Western Electric Model #23b Speech Input Console equipped with the necessary gain controls, etc. The output of the Console was in turn coupled to the input stage of a Model #PT 6A Magnecorder tape recorder. The recording tape used for this study was 1/4 in. plastic backed 3M, #111a High Fidelity magnetic tape. The tape speed for all recordings was at the rate of 7-1/2 in. per second. With the above equipment the frequency response on the recordings was well within the distortion ratio required for high fidelity speech recording. The loudspeaker used for playback purposes was a Western Electric, Model #728b, fifteen inch speaker, with an approved distortion ratio in the frequency range from 30 to 15,000 cycles per second.

See Figure 1. on the following page for a plan of the studio, equipment, arrangement, etc.

2.07 <u>Preparation of Speakers for Recording</u>

Each speaker was brought into the recording room and

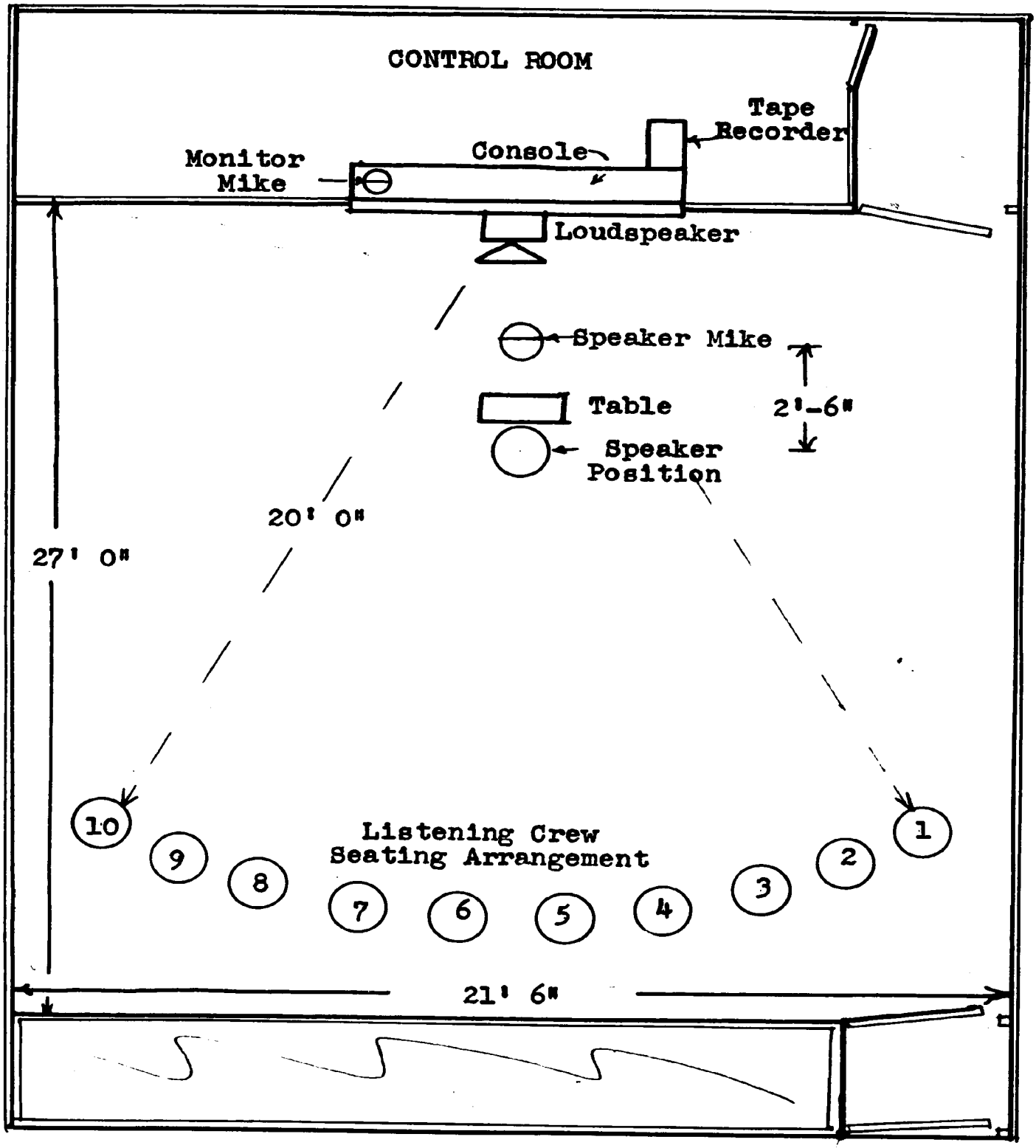

Fig. 1.—Floor Plan of Studio "C"

was seated in the chair facing the microphone. The microphone was adjusted so that it was 2-1/2 ft. from the speaker's lips. The speaker was given the selected PBF list and asked to go over the entire list out loud before the recording to correct any errors of mispronunciation. This was perhaps a moot point in the study, but the tester's judgments of the correct utterance of the word was considered as adequate. No effort was made to correct any errors of articulation of speakers in categories III and IV, of course. Some speakers occasionally "misperceived" the word by reversing letters, etc., and an attempt was made to correct such "misperceptions" during this orientation period. Then the speaker was instructed to precede each test word with the carrier phrase, "The first (or next) word is.....". As a final instruction, the speaker was requested to speak as conversationally as possible, to speak at the same loudness for each sentence, and to utter each sentence at the same rate of speed. The most difficult part of the orientation procedure was to prevent the subject from uttering the carrier phrase as a separate unit from the test word and/or from uttering the test word more loudly than the carrier phrase. The object of the inclusion of the carrier phrase was to make the complete utterance a sentence, and by doing so to eliminate the common tendency to "over-articulate" the test word. This procedure was closer to connected speech than the one used in the PB records, where the speaker utters a sequential number first, followed by a short silence

interval, and then the utterance of the test word.

The tester then went into the control room and, using the control room hand microphone, instructed the speaker to practice several speech units for monitoring purposes. The subject was instructed to watch for a hand signal from the tester, which signified that he was to utter the sentence. The tester warned him that hand signals would occur at five second intervals, and that he was to utter the sentence only upon that signal.

While the speaker was uttering these practice sentences, the studio engineer was putting a 1000 cycle test tone from the transmitter on the tape. The tone was monitored on the Console at (0) VU and plus or minus 0.25 db. on the recorder. This monitoring tone was followed by a series of short 1000 cycle tones which identified the speaker sequence for that particular tape. This procedure was necessary for check-monitoring the play back and for tape splicing purposes. Unfortunately subjects could not be recorded in the sequence necessary for the listening sessions.

2.08 Recording the Speaker

After the practice session, the recording was made using the word "word" in the carrier phrase as the monitoring peak level. This peak was attenuated to register a minus 4-1/2 VU on the Console. Generally speaking, very little monitoring was necessary for the majority of the individuals.

If the speaker skipped a word, or "mispronounced" a

word during the recording, those errors were covered by having him utter those words after the fiftieth word. (Of course, the speaker's score was not penalized for such errors). The greatest number of errors thus noted was only four for any one subject. The speakers' utterances were picked up by the control room loud speaker, so that the tester could check the deviations as the recording proceeded. A difficult aspect of this procedure was the accurate control of timing, as the reaction time for all individuals varied somewhat. This aspect was not considered as a variable as the spacing of utterances was for the convenience of the listening crew rather than for the speaker.

The recordings were taken at intervals during the months of July, August, September, and October, 1951. It would have been a great advantage to have taken those recordings in a much shorter space of time, but due to subject, engineer, and studio availability, this procedure was necessarily prolonged.

2.09 Preparation of the Recordings

As the recordings were not made according to the basic design of the experiment, the eleven tapes required had to be spliced in the design sequence and checked for sequence in preparation for the listeners. This was accomplished by playing each recording back on an Ekotape recorder, cutting out, splicing, resplicing, etc. As a part of the preparation, individual listener forms were made up and mimeographed. These forms (see the Appendix for a reproduc-

tion of Form E-1) contained the information necessary for the data required. They were numbered according to the speaker sequence, panel number, listener number, and stapled in units of sixteen sheets--one unit for each listener.

2.10 Preparation of the Studio for Listening Crews

In order to equate the actual input signal from the microphone to the tape with the output through the loud speaker, the following procedure was devised.

The microphone of a Sound Level Meter was placed adjacent to the studio microphone. The studio microphone was placed the same distance away from the speaker's mouth as it was during the recording. A studio employee went through the word list at the same level of speaking as the recorded speaker did. A reading was taken from the sound level meter during the speaking of the word, "word" in the carrier phrase. The sound level meter was then placed the same distance away from the loud speaker (2-1/2 ft.). A sample of taped material was run through the speaker. When the reading on the sound level meter coincided with the level obtained during the recording, it was felt that a fairly accurate monitoring level was achieved at the control room monitoring stage.

Desk type chairs were then arranged across the room in a slight arc facing the loud speaker--located next to the far wall (see Fig. 1.). A trial listening crew was engaged to set a level that would yield high articulation scores, (90% to 100%) when the crew listened to a normal speaker.

The following level was set. The 1000 cycle tone on the recorder ampere meter was set at a minus 10 VU. The output was fed to the speaker and bridged by the Console. The Console meter was set at (0) VU. The level of the recorder was then set so that the word "word" peaked at a minus 7VU on the Console meter. This level was approximately equivalent to a minus 17 VU on the Magnecorder.

The sound level meter was then used to check chair position so that every listener would have an equal chance to hear the recording. Fortunately the acoustic properties of the room were such that no distinguishing differences were evident at a distance of twenty feet from the loudspeaker. At this distance there was approximately a 15 db. drop in intensity from the 2-1/2 ft. mark in front of the speaker. This drop was well within the intensity range for which there is no appreciable lowering of articulation scores. The sound level of the room during a period of no speech was so low that the meter did not register at all when set to the lowest setting (30-50 db.) of the weighting network of the meter.

2.11 Selection of Listeners

Listeners were solicited on a volunteer basis and secured from students enrolled in undergraduate speech classes. All forty listeners, twenty females and twenty males, passed a pure tone audiometric sweep-check test on an ADC audiometer at 15 db. of attenuation above threshold on discrete octave frequencies from 256 to 8192 cycles.

2.12 Presentation of Recordings to the Listeners

Immediately preceding each play-back session, listener tabulation forms were passed out to each listener. Extra pencils were given out for practical reasons. The tester then read the following instructions to the crews:

> "Each speaker will read off a series of approximately fifty words at regulated intervals. Each word will be preceded by the carrier phrase, "The first (or next) word is....." That phrase is your cue to listen for the test word. Write down that word only. If you hear only a part of the word, write down the part that you hear. Try not to leave any blank spaces. Put down what you hear and not especially what you think the word should be. After the speaker has pronounced his last word, I will give you a terminal signal.
>
> Now look at your sheet and you will find a series of numbers from one to seven. Encircle one of the numbers only, which will represent your judgment of that particular speaker's intelligibility rating. This task is not a test of hearing, but of the speaker's ability to make himself "understood." Do not feel that numbers one or seven are the only ratings you can assign.
>
> Do not assign an intelligibility rating until the speaker has read his entire list.
>
> Please try to be as quiet as you can during each speaker's utterances, as each speaker is to have an equal chance to be heard undisturbed."

The tester reiterated that the test was not a test of the listener, but of the speaker, and then gave the crew an opportunity to ask any questions, etc.

After a short pause the tester gave the engineer the

the signal for the beginning of the recording. After the last word was uttered, the tester instructed the panel to assign the intelligibility rating. This period generally took about 30 seconds. After the fourth recording, the tester allowed the listeners a three minute break, after the eighth speaker--a ten to fifteen minute break to leave the studio, and another three minute break after the twelfth recording. The total running time per session was approximately two hours. All four sessions were run off in the preceding manner.

It was of some interest to note that none of the listeners was able to identify the geographical origin of the speakers of category III. Their guesses ranged from the Balkans to Hawaii. They apparently knew that some words occurred more than once, and that certain word combinations were somewhat predictable, but they were unable to remember those words at the time. The whole situation was apparently a challenge to the majority of the crews and they actively resented any unpremeditated movement or noise by any member of the crew.

2.13 Scoring the Listener Data Sheets

All words written down by the listeners were treated as errors if they did not coincide with the test words. Misspellings were carefully considered and if they deviated so far from the correct spelling that the phonetic content of the word was altered, the words were considered as errors. If the word uttered was the plural form of the word,

all singular forms were considered as errors. If the word had been "mispronounced" by the speaker and was repeated, the listener was given credit for one correct word--if his spelling coincided with the repeated word.

Articulation scores were not computed for each listener, although this factor did come under consideration in the computation of average group percentage intelligibility scores.

CHAPTER 3

PRESENTATION AND ANALYSIS OF DATA

The data derived by means of procedures outlined in the previous chapter were first analyzed in order to answer the principal problem questions.[1] Further analyses were then made in terms of several unexpected aspects of the data, which gave promise of additional significant conclusions.

3.1 Relationships between CA Scores and CA Ratings

The experimenter was first concerned with possible relationships between the two techniques of error notation and ratings by the expert examiner. To explore this question, correlations were secured for both groups of subjects tested by the examiners. (See Table 10 in the Appendix.)

When the CA Scores and the CA Ratings for the sixteen Dyslalic subjects were correlated, the r value was $0.769 \pm .105$. When the CA Scores and the CA Ratings for the other group, the Foreign Accent category, were correlated, the r value was $0.620 \pm .159$.

In interpreting the r values immediately above, the Wallace and Snedecor Tables[2] were consulted to establish minimum levels of significance. It was found that the r for category III was just barely short of the .01 level of sig-

---

[1] Cf, ante, p. 14

[2] J.P. Guilford, Fundamental Statistics in Psychology and Education, (2d ed.; New York: McGraw-Hill, 1950), 609-

nificance (.623), while the r for the Dyslalic group was substantially below the .01 level. Because of the small number of cases, it was recognized that the probable exact amount of this relationship, in the event the experiments were repeated, could only be predicted with caution, as the size of the S.E. suggested.

3.2 Relationships between the EI Scores and Ratings

The experimenter was also interested in possible relationships between the two techniques of error notation and rating by the listener crews. (Table 11 in the Appendix shows the EI Scores and Ratings for all of the listeners in all of the sessions.) Correlations were secured for all groups of subjects tested in the experimental situation and are shown below in Table 2.

TABLE 2

EI Score and Rating Correlations For The Four Experimental Speech Categories

Category	r Value	Significance Level (%)
I	$-.601 \pm .165$	Below .05
II	$-.603 \pm .164$	Below .05
III	$-.695 \pm .133$	Below .01
IV	$-.788 \pm .098$	Below .01

From these data it may be said that the correlations between variables on the first two categories suggest some fair positive relationships, but, as the size of the S.E. is

fairly large, only indicate trends in the direction of agreement. However, listeners seemed to be more consistent when they evaluated the clinical categories III and IV. It was especially interesting to note that the correlation for the Foreign Accent group was less significant than for the Dyslalic group--for which fewer errors were noted. These data and the correlations explained in Section 3.1 suggest that both listeners and expert examiners were more consistent when judging the Dyslalic group of speakers. It is also possible that there may be some rough relationship between the number of errors perceived by listeners and the number of phonetic errors noted by the expert examiners. However, the size of the sample restricts and interpretation of the efficiency of the comparison.

3.3 Relationships between the Diagnostic Test and the Experimental Test

The experimenter wondered if the judgments by the expert examiners and the untrained listeners were alike or different, and to what extent. To answer these questions, correlations were computed between the error notations (CA Scores and EI Scores), and ratings (CA Ratings and EI Ratings) for the clinical speaker categories III and IV.

Table 3 on the following page lists the relationships between scores assigned by the expert examiners and untrained listeners to performances of subjects with a foreign accent. The scores obtained from experts and listeners correlated significantly only with respect to a specific tabulation of errors. The remaining correlations,

TABLE 3

CORRELATIONS BETWEEN THE TWO TESTING TECHNIQUES IN BOTH TEST SITUATIONS FOR CATEGORY III

Description	r Value	Significance (%)
CA-S and EI-S	.728 ≠ .070	Below .01
CA-S and EI-R	-.495 ≠ .195	Above .05
CA-R and EI-R	-.410 ≠ .215	Above .05
CA-R and EI-S	.259 ≠ .241	Above .05

between ratings and tabulations, are not significant. It is apparent that experts and untrained listeners rated different variables in evaluating the intelligibility of the Foreign Accent group (Category III). That listeners and experts error tabulations were so closely related is important. This method of assessment may be the only one to secure a valid and reliable accounting of speech aberrations for this particular type of speaker.

Since the tabulation of specific errors was not available to the untrained listeners, it is possible that they may have been misled in their ratings by their overall impressions of the intelligibility of the speech samples.

As indicated by Table 4 on the following page, the results for the Dyslalics are quite different from those of the Foreign Accent group. Not only are the clinical articulation scores (CA-S) and experimental intelligibility scores (EI-S) significantly related for the Dyslalics, but

TABLE 4

CORRELATIONS BETWEEN THE TWO TESTING TECHNIQUES IN BOTH TEST SITUATIONS FOR CATEGORY IV

Description	r Value	Significance (%)
CA-S and EI-S	.602 ≠ .103	Below .05
CA-S and EI-R	-.471 ≠ .201	Above .05
CA-R and EI-R	-.567 ≠ .170	Below .05
CA-R and EI-S	.502 ≠ .193	Below .05

all other correlations are significant at approximately the five per cent level or less. (The six per cent chance probability of the difference between CA-S and EI-R is close enough to significance to be interpreted in that light in the present discussion)

How is the greater agreement for the Dyslalics than the Foreign language group to be explained? A number of possibilities may be considered. Since raters spoke the same language as the Dyslalics, the frame of reference of the listener was thus more familiar than that provided by a foreign dialect. A second factor may have been that Dyslalics had predominately fricative and glide defects. An inspection of Table 13 in the Appendix show that the errors of the Foreign Accent group contain a much larger number of consonants and vowels than were tabulated for the Dyslalics.

The results contained in Table 4 suggest a relationship between intelligibility (EI-S and EI-R) and the severity

of articulatory defects (CA-R) for Dyslalics. The subjective, relatively rapid process of rating is significantly related to and approximately as reliable as the tedious chore of tabulating errors for this category of subjects. From these data there is no basis for judging whether the subjective or objective method is superior for measuring intelligibility.

An inspection of individual intelligibility and severity ratings shows that the untrained listeners had a tendency to judge Dyslalics slightly more harshly than did the experts. This tends to contradict a popular argument that specialists in speech are hypercritical of persons with this particular type of speech defect.

Judgments of Indian speakers by untrained raters were even more severe than they were for the Dyslalics. It must be kept in mind however, that the untrained listeners did not see the speakers and heard recordings of their voices. Furthermore, they were listening to highly controlled and restricted samples of speech. The experts, on the other hand, saw each speaker and also heard a sample of connected discourse. It therefore seems probable that the job was more difficult for the untrained listeners.

In view of the different experimental situations, it is thus not possible to conclude that the untrained raters judged more severly than the experts, but the partial evidence in favor of this hypothesis warrants further research. This fact should be kept in mind and corrected subjectively by the practicing clinician who uses such ratings.

The preceding discussion indicates that ratings of the intelligibility of American Dyslalic speakers by either expert examiners or untrained listeners are reliable, while ratings of speakers with this particular type of foreign accent are not. The method of error tabulations may be the only one to secure a valid and reliable accounting of articulatory aberrations for the speaker who possesses this particular kind of foreign accent.

3.4 Differences between the Four Speaker Categories

The experimenter next considered the principal question which it was hoped the experiment would answer. Would any of the several combinations of judges and tests differentiate the four chosen categories of speakers?

To explore the problem, two separate sets of data were available. One set was concerned with the clinical situations. Since information was available on the clinical population, i.e., those subjects who had exhibited demonstrable articulatory errors, categories III and IV were analyzed. The other data were concerned with the experimental situation. The design made it possible to secure information on all four speech categories.

3.41 The Diagnostic Test Situation

In the clinical situation, the number of errors served to answer one aspect of the question. A tabulation of errors was made of categories III and IV. Table 13 in the Appendix shows a complete summary of errors noted in both categories by the expert examiner.

Page 43 missing in numbering when material checked for microfilming.

UNIVERSITY MICROFILMS

An inspection of the error totals showed that category III showed more than twice as many errors as category IV. A *t* ratio, computed from the means of the severity ratings of the two categories showed the very significant difference of 3.5 which was sufficient to reject the null hypothesis that the two groups were taken from the same population--the chances were less than one in a hundred that the two groups were not positively related. The combination of the two types of evaluations tended to strengthen the hypothesis that the number of errors and the rating scale techniques, when combined, were adequate measurements of differentiation.

3.42 The Experimental Test Situation

The Graeco-latin Square design, as was previously explained in Section 2.022, was chosen because it seemed peculiarly well adapted to answering the question of group evaluation in the experimental situation. It was assumed that due to the manner of selection, the speakers represented a stratified population and that the selection of the listeners was basically randomized. The problem therefore became one of testing the speakers and not of listener consistency in the four listening sessions. The application of the analysis of variance[1] was to test the over all significance of the differences among the four stratified factors, and at the same time to test to what extent the re-

[1] Q. McNemar, Psychological Statistics, (New York: John Wiley & Sons, 1949), 289-307.

maining variables were actually randomized. Tables 5 and 6 show the derivation of the F or Variance Ratio for both EI Scores and the EI Ratings.

3.421 Speaker Category Variance

Both analyses showed rather outstanding evidence to support the assumption that speaker categories were differentiated by both techniques of evaluation, as demonstrated by the extremely high F ratios. To further test those differences, t ratios were computed between the scores of the two categories which presented the smallest of interval differences. In the instance of the EI Scores, the smallest difference occurred between the Normal speakers and the Slovenly speakers. When subjected to the t test of significance, the difference was expressed by a t score of 2.86, which was significant below the .01 level. In the instance of the EI Ratings, the smallest difference occurred between the Slovenly speakers and the Dyslalic group. The t score expressed by those groups was 4.10--or significant below the .001 level. Both scores definitely showed that differences were very significant. It seemed safe therefore to assume that even if minimal differences were the only basis for consideration, the null hypothesis could be rejected with the confidence of being wrong at least only once in 100 times.

From the raw data, another method might show the placing of these four categories on an Intelligibility continuum. The error scores and the rating scores were conver-

TABLE 5

ANALYSIS OF VARIANCE FOR THE EI SCORES

Source	S.S.	df	Var.Est.	F	Significance-%
Rows	54,011	3	18,003.7	1.22	Above.05
Columns	147,667	3	49,222.3	3.35	Above.01
Categories	22,472,419	3	7,157,473.0	486.99	Below.001
Word Lists	510,891	3	170,297.00	11.59	Below.001
Sex	2,601	1	2,601.0	.18	Above.05
Remainder	9,200,522	626	14,697.3		
Total	31,399,111	639			

TABLE 6

ANALYSIS OF VARIANCE FOR THE EI RATINGS

Source	S.S.	df	Var.Est.	F	Significance-%
Rows	12,563	3	4,187.7	4.15	Above.001
Columns	2,939	3	979.7	.97	Above.05
Categories	632,379	3	210,793.0	209.03	Below.001
Word Lists	2,987	3	995.7	.99	Above.05
Sex	11,449	1	11,449.0	11.35	Below.001
Remainder	631,282	626	1,008.4		
Total	1,293,599	639			

ted into percentage scores, and the two scores averaged to show a smoothed "Intelligibility Score." The average number of errors for each category was multiplied by two and subtracted from 100; and the average rating scores were divided by six. (The rating scale, though it had seven points, had only six intervals.) The percentage scores, according to their order of "Intelligibility" are tabulated below.

TABLE 7

AVERAGED EI SCORES AND EI RATINGS
CONVERTED TO INTELLIGIBILITY SCORES
AND INTELLIGIBILITY SCORES IN RANK ORDER

Order	Category	Intelligibility Score	Intelligibility Rating	Combined Average Score
1.	I	91 (%)	98 (%)	94.5 (%)
2.	II	88 (%)	87 (%)	87.5 (%)
3.	III	83 (%)	77 (%)	80.0 (%)
4.	IV	54 (%)	43 (%)	48.5 (%)

From these data it was apparent that not only were there differences between categories, but that all techniques showed trends in the same direction. Apparently the naive listeners tended to downgrade the clinical categories by the ratings they assigned those speakers. The dispersion of intelligibility scores from the combined average in the two clinical groups may suggest that with an increase in the number of errors, listeners tended to assign more severe ratings--not always consistent with their perception of words.

3.5 Sex Variance

Although the evaluation of sex differences was a problem under consideration, its importance as an important variable was minimized by the design of the experiment, i.e., by randomly assigning both males and females as listeners to all of the speakers, any variance derived would be unbiased by the sex of the listener. Evidently, from the extremely low F, computed from the EI Scores, both male and female listeners were able to perceive the same number of errors. A t test between the ratings assigned by males and females showed that the female listeners rated the male speakers significantly higher than did the male listeners--as evidenced by the high t of 4.77, which was well below the .001 significance level. This was interesting in so far that it was apparently contradictory to an assumption derived from a review of the literature, that women judges tended to rate men more severely than do men judges.[1]

3.6 Word List Variance

Although it was assumed that the word lists were fairly equated to phonetic content and relative difficulty, it was apparent from the analysis of the EI Scores that significant differences did exist. An inspection of the raw scores showed that listeners noted fewer errors in List "C," or PBF 3.

Two measures of significance were taken, one between the smallest difference between Lists "A" and "C,"

---

[1] Cf, ante, p. 19.

and another between the largest difference shown by the data (List "B" vs. List "C"). These ratios, expressed by t scores of 2.43 (above .01) and 5.34 (below .001) respectively, showed that definite differences were exhibited. The fact that the EI Scores in three session were consistently lower than the scores for the other three lists gave a fairly accurate picture of the list itself. Why the list was apparently "easier" was not demonstrable by the phonetic content of the list. It was possible, of course, that the speakers uttering that list were more "intelligible" as a group than the speakers who uttered the other lists. The very low F on the EI Ratings (.99) tended to support the contention that the speaker category variable minimized the effects of any concommitant factor.

3.7 Row and Column Variance

The analysis by Rows and Columns for both testing techniques indicated by the controls intended in the design were not completely effective. By the very nature of the Graeco-latin Squares, the Rows became the speaker sequence factor and the columns became a check on the internal consistency of the design. The very low F of the EI Scores, suggested that listeners were not affected by the speaker sequence throughout the experiment as far as their (the listeners') ability to perceive errors was concerned. The listeners were obviously affected by the sequence when they estimated the speakers' EI Rating, as expressed by the very significant F of 4.15. Evidently the listeners tended to judge the

speakers more severely as each session progressed.

The significant F of 3.35 expressed by the EI Scores in the Column treatment, suggests that some learning may have taken place. An inspection of the columns by sessions in Table 14.1 in the Appendix shows that the EI Scores tended to become larger during the progress of all of the sessions. This factor was not significant in the EI Ratings, however, and led to the conclusion that, despite the fact that listeners perceived more errors, their evaluation of the speakers' total "Intelligibility" was relatively unaffected. The analysis did not reveal however, whether this Row and Column interaction was a function of the listeners or of the speakers.

3.8 Summary of the Interpretation of the Data

The data derived from the clinical and the experimental situations have been presented and analyzed. Although the correlations between techniques in both situations were based upon a relatively small number of cases and had a proportionally high standard error, they did indicate that some positive relationships existed between the two different approaches to the problem of clinical and non-clinical evaluations of persons with defective articulation. In the experimental situation the high F ratios for both evaluation techniques showed that it was quite unlikely that any of the four speaker groups contained a statistically significant number of members of the other populations.

CHAPTER 4

A DISCUSSION OF THE EFFECTS OF PHONETIC ERROR UPON THE PERCEPTION OF SPEECH

4.1 The Overview

Although Section 3.3 presented correlations between the two testing techniques in the Clinical and the Experimental test situations, those relationships were based upon the number of word errors in the latter and upon phonetic errors in the former situation. It was felt that perhaps there might be some relationship between the kinds of phonetic errors noted by expert examiners and the kinds of errors written down by the listeners. These qualitative aspects of the CA Scores and the EI Scores were defined as follows:

4.11 Clinical Articulation Error Scores (CAE Scores) represented the number and type of phonetic errors noted by expert examiners.

4.12 Experimental Sound Error Scores (ESE Scores) represented the number and type of sound errors within the words misperceived by the naive listeners.

4.2 Selection of a Population Sample

To explore possible phonetic relationships between the two test situations, a random sampling of sixteen speakers was taken from the sixty-four so that speech categories were represented equally and lists were represented by each speech category. The random sample also included

listener data from all four listener sessions. To be assured that the sample was representative of the entire speaker population, average EI Scores were converted into intelligibility scores and compared with the scores for all sixty-four speakers. With the possible exception of category IV, which was eight per cent lower than the total average intelligibility scores for that category, the other three scores were within two per cent of the other totals. To check another aspect of randomness, EI Scores were inspected according to word lists and showed that PBF List "C" had the fewest number of noted errors - which was typical of the original sampling.

4.3 Preparation of Material for Discussion

The one hundred sixty listener data sheets were analyzed for "sound errors" as written down by the listeners. The experimenter took all variations of spelling and writing into consideration and recorded only those errors which indicated that the listeners had clearly misperceived the stimulus word. These raw data were combined into Table 15 (see Appendix) which presents the tabulation of errors by speakers and by sounds. The totals at the foot of each column represent errors per ten listeners. EI scores for each speaker are listed below those ESE total scores.

Table 7 shows the averaged distribution of sounds contained in the four PBF lists and the averaged ESE scores for the sample according to sound categories. Averaged EI scores are shown at the foot of the speech category columns.

TABLE 7

AVERAGED PHONETIC CONTENT OF ALL PBF LISTS
AND
AVERAGED ESE AND EI SCORES -ALL SPEECH CATEGORIES

Sound Category	PBF Content	PBF Percent	Cat.I (ESE)	Cat.II (ESE)	Cat.III (ESE)	Cat.IV (ESE)
Plosives	33	23.2	1.5	2.0	9.6	5.1
Fricatives	25	17.6	2.0	2.4	8.3	2.7
Nasals	11	7.7	.4	1.0	2.6	1.4
Affricates	3	2.1	0.0	.2	1.2	1.6
Glides	19	13.4	.9	.8	5.6	2.2
Vowels	33	25.2	.5	1.1	6.8	2.1
Diphthongs	13	9.2	.3	.2	2.9	.4
Neutral Vowels	5	3.6	.3	.4	1.9	.3
Totals	142	100.0	5.9	8.1	38.9	15.8
EI Score Totals			4.8	7.1	23.8	12.3

4.4 Discussion of EI and ESE Scores

The total ESE scores by speech category bore a marked resemblance to the averaged EI scores of their respective categories, and in the same rank order of frequency of error. In each case, the ESE score was slightly above the corresponding EI score. Generally speaking, speakers tended to make slightly more than one error per word. Of the four types of speakers, those from the clinical population tended to make slightly more errors per word than the Normal and the Slovenly group.

The ESE score differences between the Normal and the Slovenly groups were so small that it was doubtful whether sound errors per-se contributed much to differences between the intelligibility of the two groups. This evidence tended to support the information noted in Section 3.421, in which it was shown that although differences between groups were considered statistically variant, the smallest difference between all speech categories was between those two groups. Table 8 did show, however, that most of the distinguishable differences between the groups were consonantal. Previous studies in intelligibility testing have indicated that consonants tended to lower the intelligibility scores of normal speakers under similar controlled conditions. This factor was taken into consideration in the discussion of speakers with defective articulation.

4.5 Discussion of Error Differences and Similarities Be-

tween Category III and Category IV

Table 9 presents the phonetic content of Form C-2 used in the clinical situation (see Appendix) and the averaged CAE scores taken from the sample of Categories III and IV.

TABLE 9

PHONETIC CONTENT OF FORM C-2
AND
AVERAGED CAE SCORES FOR CATEGORIES III AND IV

Sound Category	C-2 Content	C-2 Percent	Cat.III CAE	Cat.IV CAE
Plosives	52	14.2	7.8	1.5
Fricatives	65	17.7	18.8	12.3
Nasals	39	10.7	3.0	.8
Glides	38	10.4	9.3	7.5
Affricates	4	1.1	.5	2.5
Vowels	91	24.9	13.0	.3
Diphthongs	32	8.7	1.5	.3
Neutral Vowels	45	12.3	1.5	0.0
Total	366	100.0	55.4	25.2

It should be noted before discussing the clinical population that any attempt to equate Form C-2 with the PBF word lists according to specific phonetic content would be rather questionable. The clinical test had twice as many consonants as the word lists and more than three times the number of vowels. It was doubtful whether two such materials so disparate in phonetic proportion and syllabic construction could be practically equated. Of much more value would be a discussion of the relative proportions of errors between sound categories for the two types of speakers in each test situation.

Although the total averaged scores of the two categories were slightly higher than those reported for the original groups shown on Table 13 in the Appendix, both scores were within fifteen per cent of the total error mean. From the total CAE scores and the ESE scores for categories III and IV, it was observed that expert examiners and naive listeners scored the two groups similarly despite the differences of a "face-to-face" versus a recorded voice situation. They also noted slightly more than twice as many errors for III than they did for IV. This proportion was even more interesting when related to the total number of

word errors noted down by listeners. In the experimental situation, the forty listeners received EI scores in exactly the same proportion between the two clinical groups of speakers. The EI scores for the sample showed a slightly lower proportion of two to one.

An examination of the sample ESE scores and EI scores on Table 8 showed that the sound-to-word proportion of III was slightly higher than that of group IV by approximately twenty-five per cent. A discussion of the sound errors within words might explain some of these differences.

From Tables 8 and 9 it was apparent that consonant-vowel proportions varied considerably between the two clinical groups in both test situations. The consonant-to-consonant relationship was fairly even, i.e., both types of listeners noted down about twice as many consonantal errors for III as they did for IV. The vowel-to-vowel relationship was much more unstable. The naive listeners detected four times as many vowel errors in the Foreign Accent group as in the Dyslalic group, while the expert examiners detected only a trace (below one error) in the latter group compared with sixteen errors in III. However, the number of vowel errors in both situations for IV were so small that any percentage comparison would be rather spurious. Perhaps a better way to express the relationship would be to say the III had a definite number of vowel errors, predominantly of the "pure" vowel type, which was demonstrated in both test situations. Vowel errors in IV were relatively

insignificant.

The relationships already discussed express a general tendency for both types of test situations to reflect the general clinical findings for the two clinical categories. Both expert examiners and naive listeners detected similar broad types of errors which tended to differentiate the two groups of speakers with articulatory difficulties. Perhaps before going on any farther, one important factor in the experimental situation should be considered. Table 8 showed the number of sound errors for Category I, which, though small, might affect the scores of the other three groups. If plosive and fricative errors were to be expected normally, it therefore followed that speakers with defective plosives and fricatives should not be overly penalized for such errors. Even when those errors were discounted, listeners misperceived plosive sounds uttered by foreign accent speakers twice more than for Dyslalic speakers. They detected three times as many fricative errors and two and a half times as many glide errors for III and IV. Apparently the naive listeners were not as sharply critical of the speakers with fricative and glide defects as were the expert examiners. There was a possibility too that the affricates with their fricative components might contribute to raise both fricative scores. Their contribution, however, would be relatively the same for both speech categories so the general effect would not be substantially changed. The nasal sound, particularly the

(n) and (n) blends, also seemed to be of definite importance in contributing towards an increase of errors in Category III.

4.6 Interpretation of the Sample Data

Despite error variables produced by the experimental design, untrained listeners did perceive articulatory errors in an experimental situation. Undoubtedly errors of articulation contributed to the lowering of intelligibility scores. It was evident from the data that with a rise in the number of word errors, sound errors increased by more than a one to one proportion. This was revealed by the comparison of the clinical and non-clinical groups of speakers. Though consonantal errors seemed to account for the greatest number of errors in both clinical categories, vowel errors contributed to approximately one third of the total error score in the Foreign Accent category.

It was also possible that the expert examiners' diagnosis of the articulatory status of the Dyslalics influenced their actual discriminatory ability. In other words, if the speaker was a "lisper," the experts may have marked automatically the (s) sound as defective--even though the sound was only inconsistently substituted. The C-2 test sentences, however, did contain twice as many glides and approximately two and a half times as many fricatives as the PBF word lists. However, untrained listeners did perceive some errors above those uttered by the Normal group and to the extent that those factors did contribute to a

lowered intelligibility score.

An interesting difference of uttered sounds by group IV was the relatively high incidence of plosive error in the experimental situation as compared with the almost negligible amount noted by the expert examiners. It was quite possible that clinicians tended to disregard plosive errors when examining lispers and Dyslalic speakers with (r) and (l) errors. This was not too surprising in view of the fact that clinicians might tend to "look for" only fricative and glide errors, despite a considerable number of chances to also detect plosive errors in a series of diagnostic test sentences.

In summary, this section did show certain areas of differences and of communality between the two widely different types of situations. Although based on a restricted sample of speakers and listeners, the tentative results might be indicative of certain trends to be expected in larger samples of similar populations.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

5.1 Summary

This study was concerned with the comparative evaluations of intelligibility by untrained listeners and expert examiners. The significance of the problem definitions of controversial terms, and a review of the literature were discussed. There were two different test situations. Speakers were first evaluated by expert examiners who prepared a phonetic analysis and then rated on a seven point rating scale indicating the severity of the articulatory defect. Subjects' utterances of selected words were then recorded on magnetic tape and the recordings were played back in a prescribed order to untrained listeners. The subjects were sixty-four male speakers divided into four groups of speakers representing equally the following types of articulatory performance: Normal, Slovenly, Foreign Accent, and Dyslalic. The stimulus materials used in the experimental situation were randomized versions of C.V. Hudgin's four PBF (Phonetically Balanced, Familiar) word lists. The data from the two test situations were analyzed statistically and interpreted. Possible relationships between phonetic errors and listeners' perceptions of words were also analyzed.

5.2 Conclusions

5.21 Error tabulations (EA-scores) and ratings of

severity of articulatory defects by expert examiners were found to be significantly correlated. The correlation for the Foreign Accent group was significant at approximately the one per cent level and substantially below the one per cent level for the Dyslalics. These data suggest that examiners tend to be more reliable judges of American speakers than of speakers with this type of foreign accent.

5.22 Accuracy of word perception (EI-Scores) and intelligibility ratings (EI-Ratings) by untrained listeners correlated significantly for the four types of speaker categories. There is some tendency for listeners to be more reliable judges of the Dyslalic group ($\underline{r}$ = .79) and the Foreign Accent group ($\underline{r}$ = .70). The correlations for the non-clinical groups, the Normals ($\underline{r}$ = .60) and Slovenly speakers ($\underline{r}$ = .60). were, however, also below the five per cent level of significance. This might indicate that the design of this study was slightly better suited to the evaluations of speakers with articulatory defects than of the non-clinical categories.

5.23 The relationships between the tabulation of errors (CA scores) and word perception (EI-Scores) by untrained listeners for the Foreign Accent group correlated significantly below the one per cent level. The remaining correlations, between both types of ratings (CA-Ratings and EI-Ratings) and between ratings and tabulations (CA-Scores vs. EI-Ratings-CA-Ratings vs. EI-Scores), were not significant. Apparently experts and untrained listeners were not rating the same variables for this group.

5.24 Correlations between both types of clinical and experimental measurements for the Dyslalic group were significant at approximately the five per cent level or below. The significant relationship between both techniques indicates that the rating procedure is a fairly valid as well as reliable and rapid method of assessing the intelligibility of Dyslalic speech.

5.25 The examiners detected twice as many sound errors for the Indian speakers as for the Dyslalic group. The difference between the ratings of these groups is also significant at less than the one per cent level. The tabulation of error totals combined with the rating scale technique were adequate measures of differentiation.

5.26 The analysis of variance reveals significant discriminations among the four experimental groups by untrained listeners in both word perception and intelligibility rating. This means that listener' scores and ratings were grouped into the categories selected by the experimenter. A t test of significance of the smallest difference between groups was below the one per cent level. Therefore, these results show that the two methods of testing intelligibility can be utilized with untrained listeners to segregate different types of speakers with articulatory defects.

5.27 The design of the experiment controlled for the possible differences between males and females. This was fortunate since it was found that the female listeners rated male speakers significantly higher (below the .001 level) than did the male listeners. This contradicts Guilford's

impression that female judges tend to rate men more severely than do male judges. The low F of .18 indicates no significant difference in tabulations of errors in word perception.

5.28 The listeners perceived significantly fewer errors (EI-Scores) when scoring speakers who read List "C." Differences between the other three lists were not significant. Either List "C" was easier, or by chance, the speakers assigned that list were more intelligible than speakers who uttered the other lists. The former interpretation seems more probable. That the ratings (EI-Ratings) by listeners were relatively not significant might be attributable to the fact that they were influenced, in addition to the word lists, by subjective impressions of the speakers.

5.29 Whereas the data and the conclusions drawn from the statistical analysis are reliable and valid, it was evident that restrictions imposed by the design only partially succeeded in controlling variables of presentation of speakers and lists. This is shown by the Row and Column variance which do indicate some change in the performance of the listeners, who tended to assign lower ratings as each session progressed.

5.3 Summary of Findings in Chapter 4

A survey was made of phonetic errors noted by expert examiners and the errors in words misperceived by listeners from a small random sample of members from the four speech categories. Although not to be construed as conclusive evidence, results showed that possibly there was a direct relationship between the number of sound errors and the misperception of words in the experimental situation. There were indications that expert examiners and untrained listeners tended to be more consistent when perceiving consonantal errors in the clinical population, although the general vowel consistency was rather surprising, considering the differences of phonetic content between the two different types of test material.

This section did show certain areas of differences and of communality between the two widely different types of situations. Although based on a restricted sample of speakers and listeners, the tentative results might be indicative of certain trends to be expected in larger samples of similar population.

5.4 Implications

From the above conclusions, it must be evident that in both clinical and experimental situations, expert exam-

iners and naive or "untrained" listeners could distinguish differences between speakers with articulation defects and those without articulation defects. It also became apparent that although no data were available in the clinical situation (by definition), teachers of speech were able to select Slovenly speakers and Normal speakers distinguishable in an experimental situation. The discussion in Chapter 4 suggested that possibly some positive relationships exist between an expert notation of phonetic errors by using clinicians and the untrained listeners' perception of the phonetic content of words. An interesting speculation surrounded the techniques of the rating scale. From the data gathered in this study, it might be presumed that the severity of the articulatory defect was positively related to rated intelligibility, especially when it came to rating the Dyslalic group of speakers. This might indicate that clinicians could predict the intelligibility of Dyslalic speakers by listening to recorded samples of speech. This procedure could be more precisely and accurately measured by experimentally testing substantially larger groups of Dyslalic speakers and using more expert examiners.

Any attempt to generalize upon the data presented was bound to be tenuous. Within the boundaries set up in this study, intelligibility was restricted to monosyllabic word identification, which, in the heirarchy of information units available to the listeners, is considered to be near the bottom of the list. Intelligibility scores are propor-

tionally high for spondaic words, and higher yet for sentences. In terms of current information theory, we are logarithmically reducing the possibility of alternates or errors when we introduce concepts of context. Whatever bases we have for the possible prediction of the intelligibility of connected discourse then becomes purely speculative. We might assume that words, being components of connected discourse, present a minimal prediction.

In this study, the terms "severity" and "intelligibility" possessed some equable meaning in a sharply restricted sense. The implication being that the two terms, despite certain definite denotations, might be aspects of the same factor and that errors might cumulatively increase the possible alternates during connected discourse. On the other hand, we have the force of context impinging upon the "gestalt" which tends to reduce the possibility for "error" in discourse. When we ask a listener to judge "severity" or "intelligibility" then, despite our efforts to delimit the meanings of those terms, we are never sure just what criteria the listener will use. In the case of an American listener who is asked to write down the words uttered by a foreign speaker, we may be confronted by an entirely different communication procedure. Despite the fact that the foreign speaker possesses some facility with the American language, the listener may be immediately prejudiced in a variety of ways, none of which may have a relationship to the hearing process at all! Perhaps a more fair way of

judging the foreign speaker would be to have a group of listeners who have the same language backgrounds judge those speakers by both of the techniques used in this experimental situation. These listeners might possibly, or even probably, assign higher intelligibility scores than did the American listeners in this study.

Of interest, too, in the communication procedure, is the question of "native" listener stratification. It is entirely conceivable that certain samples taken from the entire population would tend to be "better" or "worse" listeners than the college students used in this study. Undoubtedly, larger samples of the population taken from various professions and skills would yield different results. This might also be said of the age of the listener, bearing in mind that all samples would have equal chances to "hear" the speaker.

As a final statement, it is this experimenter's opinion that the experimental techniques demonstrated in this study might be much more practical for the evaluation of the articulatory problems of the young child. If a method of presentation could be contrived to evaluate the intelligibility of the non-reading Dyslalic child, the field of speech correction would be able to contribute answers to two very important questions in the study of intelligibility:

1) What is the intelligibility status of the child with an articulatory disorder?

2) Could various measures of intelligibility be made during the period of therapy to show changes in articulatory efficiency?

APPENDIX

TABLE 10

PHONETIC CONTENT OF FORM C-2 AND THE FOUR PBF LISTS
PART I - CONSONANTS

Sound	Clinical Test		PBF-A		PBF-B		PBF-C		PBF-D	
p	17		6		4		5		6	
b	9		6		9		1		6	
t	10		9		11		6		8	
d	6		6		7		7		4	
k	5		0		2		9		9	
g	5		4		1		4		2	
(blends)										
kw-	0	52	0	31	1	35	0	32	0	35
f	3		3		2		4		4	
v	5		2		2		2		1	
θ	2		0		0		2		1	
ð	15		3		2		0		0	
s	11		5		5		4		6	
z	10		4		2		4		1	
ʃ	5		3		4		3		1	
ʒ	2		0		1		0		0	
h	7		4		3		2		3	
(blends)										
-ft	0		1		0		0		0	
-s-	5	65	5	30	2	23	4	25	7	23
r	10		9		2		1		4	
l	10		0		3		5		0	
L	1		2		1		1		1	
j	3		0		0		0		1	
w	6		0		2		1		0	
ʍ	3		1		0		2		0	
(blends)										
-l-	1		1		4		4		3	
-r-	4	38	10	23	7	19	8	22	4	13
m	10		0		2		4		5	
n	15		6		5		5		6	
ŋ	6		0		2		0		0	
(blends)										
-mp	0		0		1		0		0	
-n-	7		2		1		0		4	
	1	39	0	8	1	12	0	9	0	15
tʃ	3		1		1		1		2	
dʒ	1	4	0	1	1	2	2	3	3	5

TABLE 10 - Continued

PART II - VOWELS

Sound	Clinical Test		PBF-A		PBF-B		PBF-C		PBF-D	
i	16		3		4		2		4	
I	21		6		5		4		6	
ɛ	18		6		4		9		5	
æ	9		7		8		7		6	
a	12		5		6		4		7	
ɔ	4		1		4		4		1	
U	4		0		0		2		2	
u	7	91	1	29	3	34	1	33	3	34
OU	11		8		2		2		2	
eI	2		0		4		4		3	
aI	1		0		0		1		1	
aU	8		2		2		7		3	
ɔɪ	9		7		1		1		2	
ju	1	32	0	17	1	10	0	15	0	11
^	32		3		5		0		4	
ɝ	13	45	1	4	1	6	3	3	3	7
Total		366		143		141		142		143

TABLE 11

CA SCORES AND RATINGS
IN THE DIAGNOSTIC ARTICULATION TEST

Category III			Category IV		
Spkr.	CA-S	Mean CA-R	Spkr.	CA-S	Mean CA-R
1.	44	3.0	1.	38	4.0
2.	66	4.0	2.	27	3.0
3.	70	4.0	3.	27	3.5
4.	48	5.0	4.	18	2.0
5.	59	4.0	5.	11	1.0
6.	68	4.0	6.	25	2.0
7.	49	4.0	7.	15	1.0
8.	71	4.5	8.	22	2.0
9.	32	1.5	9.	17	2.0
10.	48	2.5	10.	23	3.0
11.	57	3.5	11.	22	3.0
12.	47	3.0	12.	17	3.0
13.	26	3.0	13.	27	3.0
14.	30	3.5	14.	28	2.0
15.	24	2.0	15.	22	2.5
16.	27	4.0	16.	8	1.5
Total	766	55.5		347	38.5
Mean	47.9	3.5		21.7	2.4

TABLE 12

EI SCORES AND RATINGS BY SESSIONS

LISTENING CREW 1

Category I

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	M	9	5	3	7	11	6	11	4
2	M	4	6	1	7	6	7	7	4
3	F	4	6	2	7	3	7	7	6
4	F	6	5	3	7	6	7	7	3
5	F	7	6	3	7	3	6	7	3
6	F	12	6	8	7	6	7	19	5
7	F	6	3	4	7	7	6	8	5
8	M	5	5	2	6	6	5	12	5
9	M	7	5	4	7	4	8	13	4
10	M	5	5	2	6	3	7	8	4
TOTAL		65	52	32	67	55	65	99	43

Category II

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	M	5	3	11	5	4	5	13	5
2	M	2	6	4	6	2	6	7	5
3	F	4	7	6	6	4	7	6	5
4	F	5	6	6	6	5	7	5	6
5	F	4	5	3	6	5	6	8	4
6	F	8	7	16	6	7	7	13	6
7	F	5	4	6	2	5	6	10	6
8	M	3	5	4	4	3	4	9	4
9	M	3	6	9	5	5	6	7	4
10	M	3	6	5	5	2	5	6	4
TOTAL		42	55	70	51	42	59	84	49

TABLE 12--Continued

LISTENING CREW 1

Category III

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	M	39	1	34	2	29	2	35	2
2	M	25	2	32	1	19	2	28	2
3	F	24	3	33	3	18	3	25	2
4	F	36	2	33	2	21	3	26	2
5	F	30	3	27	4	25	2	27	2
6	F	38	3	38	2	24	4	32	3
7	F	34	2	38	2	19	3	30	3
8	M	32	2	33	2	20	2	33	1
9	M	30	2	33	2	17	2	30	2
10	M	29	2	29	2	18	3	29	2
TOTAL		317	22	330	22	210	26	295	21

Category IV

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	M	15	4	4	6	16	5	9	5
2	M	11	4	4	5	7	5	8	5
3	F	10	7	4	6	5	5	8	5
4	F	13	6	2	7	2	6	7	5
5	F	8	5	3	5	4	4	7	4
6	F	17	6	4	7	14	7	15	6
7	F	11	5	6	7	9	5	9	5
8	M	13	4	5	4	10	4	6	4
9	M	15	3	4	6	7	5	9	5
10	M	9	6	2	4	4	5	7	6
TOTAL		122	50	38	57	78	51	85	50

TABLE 12--Continued

LISTENING CREW 2

Category I

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	7	7	8	6	6	7	3	7
2	F	3	7	2	6	5	7	5	5
3	F	1	7	4	5	7	5	4	4
4	M	1	7	6	5	8	5	6	7
5	M	1	7	1	4	4	5	5	2
6	M	1	7	3	6	8	4	7	4
7	F	1	7	3	6	10	5	5	4
8	M	1	7	4	6	6	6	5	6
9	M	4	6	1	5	11	5	11	4
10	F	2	7	0	7	5	7	2	6
TOTAL		22	69	32	56	70	56	53	49

Category II

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	9	4	4	7	4	5	10	4
2	F	9	5	1	7	9	6	11	4
3	F	7	5	4	7	5	2	8	4
4	M	7	6	1	6	6	5	13	3
5	M	8	3	5	4	5	2	6	1
6	M	14	6	2	4	9	5	16	2
7	F	9	3	4	5	6	5	11	3
8	M	6	4	3	4	4	4	6	2
9	M	15	4	6	6	8	5	24	1
10	F	4	6	3	6	5	15	11	3
TOTAL		88	46	33	56	61	44	116	27

TABLE 12 --Continued

LISTENING CREW 2

Category III

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	22	3	19	2	20	3	27	2
2	F	14	4	13	4	20	3	21	2
3	F	18	5	13	4	16	5	21	3
4	M	18	4	15	5	22	3	24	2
5	M	13	3	14	2	20	1	23	1
6	M	13	3	16	2	24	1	22	2
7	F	15	3	16	3	21	2	18	2
8	M	16	2	10	3	20	2	21	1
9	M	19	3	19	3	26	2	27	1
10	F	14	4	13	5	15	4	17	3
TOTAL		162	34	148	33	204	26	221	19

Category IV

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	15	6	1	5	10	4	9	5
2	F	7	6	1	7	2	5	9	6
3	F	7	6	1	6	8	3	9	4
4	M	10	5	0	6	12	4	15	4
5	M	6	6	2	3	10	3	9	1
6	M	7	5	0	7	7	2	11	3
7	F	5	5	1	6	6	4	12	3
8	M	6	6	1	6	7	4	6	3
9	M	11	3	2	5	8	4	16	4
10	F	8	6	1	7	5	5	6	5
TOTAL		82	54	10	58	75	38	102	38

TABLE 12 --Continued

LISTENING CREW 3

Category I

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	9	7	7	7	2	6	5	7
2	F	1	6	7	4	1	6	7	1
3	M	4	6	4	6	2	7	1	7
4	M	5	6	5	5	1	7	1	7
5	F	3	7	7	7	0	7	1	7
6	M	5	7	4	7	2	7	1	7
7	M	5	7	6	7	4	7	3	7
8	F	3	7	11	4	1	7	4	7
9	F	3	7	9	4	1	6	3	6
10	M	2	5	7	5	1	6	1	7
Total		40	65	67	56	15	66	27	63

Category II

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	RI-S	EI-R	EI-S	EI-R
1	F	14	6	9	5	20	2	8	3
2	F	15	6	6	5	9	4	5	4
3	M	12	5	4	6	8	3	2	5
4	M	11	6	5	7	12	5	5	7
5	F	12	6	7	7	8	4	4	7
6	M	12	6	4	6	3	6	3	6
7	M	14	7	9	7	12	4	5	6
8	F	13	6	3	6	9	5	4	6
9	F	9	6	7	4	9	6	4	5
10	M	7	5	5	6	6	4	2	6
Total		119	59	59	59	96	43	42	55

TABLE 12 --Continued

LISTENING CREW 3

Category III

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	34	5	28	2	22	1	16	3
2	F	25	2	25	3	15	4	11	4
3	M	29	1	26	3	14	4	11	5
4	M	28	2	26	4	15	5	15	2
5	F	25	2	22	3	18	3	13	5
6	M	27	2	27	4	19	5	16	5
7	M	33	2	29	5	19	5	18	4
8	F	24	2	27	1	16	4	16	2
9	F	24	2	26	1	20	4	19	2
10	M	20	3	27	4	12	3	10	5
Total		269	23	263	30	170	38	145	37

Category IV

LISTENER	SEX	Speaker 1		2		3		4	
		EI*S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	9	6	20	4	18	1	13	3
2	F	4	5	14	4	11	2	8	4
3	M	9	6	14	4	14	3	10	5
4	M	6	7	15	3	13	3	7	6
5	F	4	5	14	6	13	6	8	5
6	M	4	6	14	4	13	4	7	6
7	M	4	7	19	5	13	2	9	5
8	F	6	6	15	3	11	3	9	5
9	F	7	5	21	2	15	3	13	6
10	M	3	6	14	5	9	3	7	4
Total		56	59	160	40	130	30	91	49

TABLE 12 --Continued

LISTENING CREW 4

Category I

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	2	5	3	6	7	6	2	7
2	F	2	5	5	5	9	6	4	7
3	F	3	5	4	6	6	6	2	7
4	M	5	4	6	6	10	6	2	7
5	M	4	4	3	4	6	7	2	7
6	M	4	4	1	5	8	7	0	7
7	M	3	5	2	6	11	6	2	7
8	F	5	5	4	5	8	6	2	7
9	M	3	4	3	5	4	6	0	7
10	F	4	6	3	7	10	6	2	7
Total		35	47	34	55	79	62	18	70

Category II

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	1	5	5	6	2	7	4	7
2	F	1	6	5	6	3	7	1	7
3	F	2	5	8	6	5	7	3	6
4	M	0	6	6	4	7	7	4	5
5	M	0	6	5	5	3	6	4	6
6	M	0	5	5	5	4	6	3	5
7	M	0	6	6	5	2	6	4	6
8	F	2	5	6	4	2	6	3	6
9	M	1	6	5	5	1	7	2	6
10	F	1	5	6	6	6	6	8	6
Total		8	55	57	52	35	65	36	60

TABLE 12 --Continued

LISTENING CREW 4

Category III

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	13	3	22	2	5	2	25	1
2	F	19	4	21	2	21	3	24	1
3	F	16	3	25	1	31	3	28	2
4	M	17	3	23	1	30	2	26	1
5	M	11	3	21	1	28	1	30	1
6	M	15	2	25	2	26	2	26	2
7	M	18	3	22	1	27	2	23	3
8	F	17	3	23	2	27	2	26	1
9	M	13	3	20	1	25	1	26	2
10	F	18	2	24	2	2	2	33	1
Total		157	29	226	15	268	20	267	15

Category IV

LISTENER	SEX	Speaker 1		2		3		4	
		EI-S	EI-R	EI-S	EI-R	EI-S	EI-R	EI-S	EI-R
1	F	1	6	18	4	7	5	5	6
2	F	2	6	19	2	9	6	6	6
3	F	3	4	16	2	2	5	6	2
4	M	6	5	17	2	3	4	5	5
5	M	3	5	14	2	9	4	8	5
6	M	2	4	20	2	5	5	7	3
7	M	0	5	20	2	3	5	2	6
8	F	0	6	20	3	3	4	4	3
9	M	2	5	17	2	5	4	5	5
10	F	1	6	17	2	5	5	4	6
Total		20	52	178	23	51	47	52	47

TABLE 13

TABULATION OF PHONETIC ERRORS NOTED BY EXPERT EXAMINERS IN CATEGORIES III & IV PART I - CONSONANTS

Sound	Cat.III	S.Totals	Total	Cat. IV	S.Total	Total
p	50			2		
b	0			0		
t	14			4		
d	4			2		
k	10			1		
g	1	79		0	9	
f	0			1		
v	19			1		
ð	46			5		
ɹ	129			8		
s	13			66		
z	24			76		
ʃ	6			14		
ʒ	13			8		
h	4			3		
(blends)						
s-	1			40		
z	0	255		9	231	
r	40			23		
l	4			18		
L	0			0		
j	4			0		
w	41			1		
ʌ	41			19		
(blends)						
r-	5			11		
l-	2	137		1	73	
m	0			0		
n	0			0		
ŋ	27			4		
(blends)						
n-	6			1		
-ŋk	2	35		0	5	
tʃ	5			13		
dʒ	1	6	512	3	16	334

TABLE 13 -_Continues

PART II -VOWELS

Sound	Cat.III	S.Totals	Total	Cat.IV	S.Totals	Total
i	5			0		
I	57			0		
ɛ	18			1		
æ	42			0		
a	33			0		
ɔ	6			1		
U	23			0		
u	2	186		1	3	
oU	4			0		
eI	24			1		
aI	1			0		
aU	0			0		
ɔI	7			0		
ju	0	36		0	1	
ʌ	17			0		
ɜ	15	32	254	9	9	13
Total			766			347

TABLE 14

SUMS OF ERRORS AND RATINGS BY SESSIONS

14.1 Design

	Errors		Ratings	
	Rows	Columns	Rows	Columns
Session 1	546	430	179	179
	470	447	197	194
	385	494	201	193
	563	593	163	174
Total	1964	1964	740	740
Session 2	354	344	203	180
	223	358	203	175
	410	416	164	159
	492	361	133	189
Total	1479	1479	703	703
Session 3	484	434	206	197
	549	412	185	198
	411	539	177	182
	305	364	204	195
Total	1749	1749	772	772
Session 4	220	362	183	177
	495	388	145	187
	433	361	194	189
	373	410	192	161
Total	1521	1521	714	714

TABLE 14 --Continued

14.2 Speech Categories

Sessions	Errors			
	Cat.I	Cat.II	Cat.III	Cat.IV
1.	251	238	1152	323
2.	177	398	735	369
3.	149	316	847	437
4.	166	136	918	301
Total	743	988	3652	1330
Sessions	**Ratings**			
	Cat.I	Cat.II	Cat.III	Cat.IV
1.	227	214	91	208
2.	320	173	112	188
3.	250	216	128	178
4.	234	232	79	169
Total	941	835	410	743

TABLE 14 -- Continued

14.3 Word Lists

Sessions	Errors			
	Word Lists			
	A	B	C	D
1.	557	542	389	476
2.	447	552	374	376
3.	314	408	323	434
4.	356	345	303	517
Total	1674	1847	1389	1803
Sessions	**Ratings**			
	Word Lists			
	A	B	C	D
1.	174	187	181	198
2.	190	173	177	163
3.	202	169	191	210
4.	187	190	187	150
Total	753	719	736	721

TABLE 14 -- Continued

14.4 Sex Differences

Males

Sessions	Errors				Ratings			
	Categories				Categories			
	I	II	III	IV	I	II	III	IV
1.	123	107	574	165	111	99	38	95
2.	94	164	382	146	108	77	46	84
3.	64	141	421	204	130	113	73	94
4.	79	62	452	153	114	113	37	79
Total	360	474	1829	668	463	402	194	352

Females

Sessions	Errors				Ratings			
	Categories				Categories			
	I	II	III	IV	I	II	III	IV
1.	128	131	578	158	116	115	53	113
2.	83	134	353	123	122	96	66	104
3.	85	175	426	233	120	103	55	84
4.	87	74	466	148	120	119	42	90
Total	383	514	1823	662	478	433	216	391

TABLE 15

SOUNDS MISSED BY LISTENERS IN THE SIXTEEN SPEAKER SAMPLE
PART I -CONSONANTS

Category I

Sound	Speaker, list, and Session Numbers					
	13-C1-2	10-B2-2	5-A2-4	4-D4-4	Total	Total
p	7	5	10	3	25	
b	1	0	11	0	12	
t	0	8	2	7	17	
d	0	1	0	0	1	
k	3	0	0	1	4	
g	0	0	0	0	0	
(kw-)	0	0	0	0	0	59
f	2	3	2	0	7	
v	7	3	0	0	10	
θ	15	1	0	3	19	
ð	0	2	0	0	2	
s	0	5	4	0	9	
z	11	4	3	2	20	
ʃ	0	0	0	0	0	
ʒ	0	0	0	0	0	
h	0	0	0	0	0	
(-ft)	0	0	0	0	0	
(-s-)	6	2	1	2	11	78
m	0	0	0	0	0	
n	0	3	0	3	6	
ŋ	0	0	0	0	0	
(-mp)	0	0	0	0	0	
(-n-)	0	3	6	0	9	
(-ŋk)	0	0	0	0	0	15
r	1	0	0	7	8	
l	0	0	0	0	0	
L	0	3	0	1	4	
j	0	0	0	0	0	
w	0	0	0	0	0	
hw	6	0	1	0	7	
(-l-)	0	7	0	0	7	
(-r-)	5	3	0	3	11	37
tʃ	0	0	0	0	0	
dʒ	0	0	0	0	0	0

TABLE 15 --Continued

PART II - VOWELS

Category I

Sound	Speaker, List, and Session Numbers					
	13-C1-2	10-B2-2	5-A2-4	4-D4-4	Total	Total
i	0	0	10	0	10	
I	1	0	0	3	4	
ɛ	0	2	0	0	2	
æ	0	0	2	0	2	
a	1	1	0	1	3	
ɔ	0	0	0	0	0	
U	0	0	0	0	0	
u	0	0	0	0	0	21
oU	0	0	0	0	0	
eI	0	0	0	0	0	
aI	0	0	0	1	1	
aU	0	7	0	2	9	
ɔɪ	0	0	0	0	0	
ju	0	0	0	0	0	10
ʌ	0	9	1	0	10	
ɝ	0	0	0	0	0	10
Total	66	72	53	39	230	230
EI Score Total	53	70	34	35	192	

TABLE 15 --Continued

PART I -CONSONANTS

Category II

Sound	Speaker, List, and Session Numbers					
	14-C1-3	5-B1-1	14-D2-2	8-A2-4	Total	Total
p	1	1	0	1	3	
b	0	4	0	3	7	
t	0	18	7	9	34	
d	1	2	10	7	20	
k	1	1	6	0	8	
g	0	3	2	2	7	
(kw-)	0	0	0	0	0	79
f	1	0	0	1	2	
v	6	1	3	3	13	
θ	12	1	0	0	13	
ð	0	1	0	2	3	
s	0	8	4	0	12	
z	3	7	5	10	25	
ʃ	0	0	0	0	0	
ʒ	0	0	0	0	0	
h	0	4	1	0	5	
(-ft)	0	1	0	7	8	
(-s-)	11	1	4	0	16	97
m	1	0	5	0	6	
n	2	2	6	3	13	
ŋ	0	0	0	0	0	
(-mp)	0	1	0	0	1	
(-n-)	0	2	12	2	16	
(-ŋk)	0	2	0	0	2	38
r	0	0	0	2	2	
l	0	0	0	0	0	
L	0	3	3	1	7	
j	0	0	0	0	0	
w	0	0	0	0	0	
(-l-)	0	2	0	0	2	
(-r-)	0	3	8	2	13	
ʌ	1	0	0	7	8	32
	0	0	1	0	1	
	5	0	1	0	6	7

TABLE 15 --Continued

PART II -VOWELS

Category II

Sound	Speaker, List, and Session Numbers					
	14-C1-3	5-B1-1	14-D2-2	8-A2-4	Total	Total
i	0	0	0	0	0	
I	0	0	9	5	14	
ɛ	1	2	12	3	18	
æ	0	5	2	0	7	
a	0	0	1	0	1	
ɔ	0	2	0	0	2	
U	0	0	1	0	1	
u	0	0	0	0	0	43
oU	0	0	0	3	3	
eI	0	0	1	0	1	
aI	0	0	0	1	1	
aU	0	1	1	0	2	
ɔɪ	0	0	0	0	0	
ju	0	0	0	0	0	7
ʌ	0	2	0	0	2	
ɝ	0	0	12	0	12	14
Total	46	80	117	74	317	317
EI Score Total	42	70	116	57	285	

TABLE 14 -Continued

PART I -CONSONANTS

Category III

Sound	Speaker, List, and Session Numbers					
	3-D3-1	1-B3-3	15-A3-2	16-C4-3	Total	Total
p	34	19	19	24	96	
b	10	32	5	0	47	
t	56	33	26	18	133	
d	10	16	14	4	44	
k	36	4	0	10	50	
g	2	2	1	6	11	
(kw-)	0	2	0	0	2	383
f	8	3	0	4	15	
v	5	7	14	16	42	
θ	9	0	0	21	30	
ð	0	15	23	0	38	
s	13	9	7	1	30	
z	8	5	34	2	49	
ʃ	1	34	0	1	36	
ʒ	0	0	0	0	0	
h	21	14	9	2	46	
(-ft)	0	0	0	0	0	
(-s-)	31	9	3	3	46	
m	13	6	0	1	20	
n	14	11	13	17	55	
ŋ	0	4	0	0	4	
(-mp)	0	6	0	0	6	
(-n-)	8	4	4	0	16	
(-ŋk)	0	1	0	0	1	102
r	12	6	25	3	46	
l	0	3	0	7	10	
L	3	10	1	1	15	
j	4	0	0	0	4	
w	0	13	0	4	17	
ʌ	0	0	10	7	17	
(-l-)	11	8	2	11	32	
(-r-)	13	34	20	14	81	222
tʃ	10	4	1	0	15	
dʒ	25	4	0	2	31	46

TABLE 15 --Continued

PART II -VOWELS

Category III

Sound	Speaker, List, and Session Numbers					
	3-D3-1	1-B3-3	15-A3-2	16-C4-3	Total	Total
i	4	12	6	1	23	
I	19	3	9	7	38	
ɛ	13	11	4	0	28	
æ	30	37	0	10	77	
a	19	28	24	5	76	
ɔ	2	15	0	2	19	
U	3	0	0	1	4	
u	2	2	4	0	8	273
oU	11	0	10	4	25	
eI	17	1	9	4	31	
aI	0	0	13	0	13	
aU	10	10	0	20	40	
ɔI	7	0	0	0	7	
ju	0	0	0	0	0	116
ʌ	13	19	0	0	32	
ɝ	23	4	14	3	44	76
Total	530	460	324	236	1550	1550
EI Score Total	317	269	221	145	952	

TABLE 15 --Continued

PART I -CONSONANTS

Category IV

Sound	Speaker, List, and Session Numbers					
	6-B2-3	11-C1-2	10-A2-1	6-D3-4	Total	Total
p	9	1	4	5	19	
b	9	0	11	0	20	
t	32	6	18	20	76	
d	0	7	6	12	25	
k	8	2	5	24	39	
g	0	8	2	9	19	
(kw-)	7	0	0	0	7	205
f	7	0	1	10	18	
v	8	7	1	0	16	
θ	0	12	0	3	15	
ð	6	0	2	0	8	
s	0	0	2	3	5	
z	4	1	2	10	17	
ʃ	1	0	0	0	1	
ʒ	0	0	0	0	0	
h	1	0	0	0	1	
(-ft)	0	0	0	0	0	
(-s-)	1	8	2	15	26	107
m	5	1	0	11	17	
n	4	1	5	0	10	
ŋ	2	0	0	0	2	
(-mp)	3	0	0	0	3	
(-n-)	1	3	8	11	23	
(-ŋk)	0	0	0	0	0	55
r	1	0	1	2	4	
l	1	0	0	0	1	
L	10	0	0	1	11	
j	0	0	0	0	0	
w	11	0	0	0	11	
ʌ	0	3	4	0	7	
(-l-)	9	2	1	5	17	
(-r-)	27	3	5	1	36	87
tʃ	0	6	0	22	28	
dʒ	0	7	0	30	37	

TABLE 15 --Continued

PART II -VOWELS

Category IV

Sound	Speaker, List, and Session Numbers					
	6-B2-3	11-C1-2	10-A2-1	6-D3-4	Total	Total
i	0	0	3	0	3	
I	5	2	0	10	17	
ɛ	0	6	3	1	10	
æ	2	1	2	1	6	
a	11	3	3	3	20	
ɔ	14	4	2	8	28	
U	0	0	0	0	0	
u	0	1	0	0	1	85
oU	1	0	2	0	3	
eI	0	1	0	0	1	
aI	0	1	1	0	2	
aU	4	3	0	1	8	
ɔɪ	0	0	0	0	0	
ju	1	0	0	0	1	15
ʌ	1	0	0	4	5	
ɜ	5	0	1	0	6	11
Total	211	100	97	222	630	630
EI Score Total	160	75	78	178	491	

Reproduction of Instruction Form C-1

INSTRUCTIONS FOR PHONETIC TRANSCRIPTION AND JUDGMENT OF THE SEVERITY OF THE ARTICULATION DEFECT

1. The standard for "correct" or adequate articulation will be considered for this purpose as "general american speech."
2. Any deviation from this standard will be considered as errors in articulation and must be recorded.
3. Errors will be categorized as substitutions, omissions and distortions of phonemes.
4. Errors of vowels and consonants must be recorded, using the following code:
   a) Substitutions: Write the sound you heard below the letter in the word for which there was the substitution.
   b) Omissions: Encircle the sound or sounds omitted.
   c) Distortions: Draw a diagonal line thru the sound or sounds so distorted in the word.
5. The following phonetic symbols will be used for the purposes of transcription: (A code or identifying word follows the bracketed symbol with the letter underlined which corresponds to the pertinent sound.)

Consonant sounds				Vowel sounds	
(p) put	(r) rat			(i) eat	
(b) bat	(m) Man			(I) it	
(t) to	(n) man			(ɛ) met	
(d) dog	(ŋ) sing			(e) hate	
(k) cat	(tʃ) chew			(æ) hat	
(g) dog	(dʒ) judge			(ɑ) father	
(f) four				(ɔ) haw	
(v) vine	Semi-vowels and Dipthongs			(o) hope	
(θ) think	(l) long	(ɔɪ) boy		(U) hub	
(ð) the	(L) battle	(eI) hate*		(u) hoot	
(s) so	(j) young	(oU) hope*		(ʌ) or (ə) but	
(z) zoo	(w) wet			(ɝ) or (ɚ) bird	
(ʃ) shoe	(ʍ) when				
(ʒ) azure	(aɪ) find				
(h) hope	(aʊ) cow				

*For more rapid transcription, the dipthongs (eI) and (oU) will be transcribed as "pure" vowels-(e) and (o).

6. Instructions for transcription procedure: Have the subject read the sentences and the word list at a conversational level stating that the words are not to be "over-articulated" but to be read as if the sentence occurred in conversation. Do not have them go to the next sentence until you have finished your transcription. Try to make your notations as unobtrusively as possible.
7. Instructions for judging the severity of the articulation defect: After the subject has finished the last sentence and the word list; and you have made your notations, encircle one of the numbers at the bottom of the sheet. The judgment will be considered as unique and as a reflection of your clinical experience with speech defectives.

BE SURE THAT IDENTIFYING DATA AT THE TOP OF THE SHEET ARE FILLED IN

REPRODUCTION OF FORM C-2

Name_____Address_____Clinician_____

Date_____

1. Sassy mice race across the ice.
   Zippers are easy to close.
2. Why whisper everywhere?
   We will win anyway
3. Think nothing but the truth.
   This leather is smooth
4. Shaving lotion will splash
   Measure the position of the garage
5. Choose a pitcher to match.
6. Find the office knife.
   Violets cover the grove.
7. Lend me a ruler and pencil.
   Mamma made some lemon jam.
8. Wrap the carrots with paper.
   New pennies shine.
9. Your young onions are yellow.
   Bring the singer along.
10. Tie a better knot.
    Daddy nodded his head
11. Keep the pocket book.
    Go get a bigger egg.
12. Papa put paper in the pipe.
    Buy baby a bib.
13. He hit his hand anyhow.

1. Hot
2. Hat
3. Heck
4. Hay
5. Hit
6. He
7. Hawk
8. Hut
9. Her
10. Hoe
11. Hook
12. Who
13. High
14. How
15. Hoist
16. Hue

Check here if subject regarded as a "normal" speaker ☐

Rating Scale(Severity of Artic. Defect)

Circle one number only

1 2 3 4 5 6 7

nearly normal — very severe

Reproduction of Form C-3

ORAL READING PARAGRAPH

Winter can be a pleasant season if one takes an active interest in outdoor sports. There are so many things to do in the winter that afford pleasure as well as exercise. For instance, there is nothing quite so exciting as a flying trip down a long hill on skiis, especially when the snow is new and fine. Even an occasional tumble in the snow can be fun, if you have a sense of humor; and the thrill of a first jump successfully complete, is hard to beat. Tobogganing is another sport that is a lot of fun when the track is smooth and full of nicely banked curves. Skating, cutter riding, and ice boating all have their advantages, depending on how great your skill and how extensive your energy. Even building snow men can be fun and there is much to be said for a good old-fashioned belly slammer on a little boy's sled.

Reproduction of Listener Form (E-1)

Name_____

Speaker #_____ Date _____, 1951 Panel #_____

1_____	21_____	41_____
2_____	22_____	42_____
3_____	23_____	43_____
4_____	24_____	44_____
5_____	25_____	45_____
6_____	26_____	46_____
7_____	27_____	47_____
8_____	28_____	48_____
9_____	29_____	49_____
10_____	30_____	50_____
11_____	31_____	51_____
12_____	32_____	52_____
13_____	33_____	53_____
14_____	34_____	54_____
15_____	35_____	55_____
16_____	36_____	56_____
17_____	37_____	57_____
18_____	38_____	58_____
19_____	39_____	59_____
20_____	40_____	60_____

Intelligibility Rating Scale
(Encircle 1 number only)

\- 1 2 3 4 5 6 7 -

very poor intelligibility excellent intelligibility

PBF-1 (LIST A)

No. 1	No. 2	No. 3	No. 4
1. range	1. soap	1. bar	1. tone
2. bad	2. smile	2. toe	2. smile
3. rag	3. rub	3. such	3. slip
4. bar	4. end	4. rat	4. pile
5. crush	5. bite	5. this	5. pants
6. crash	6. ford	6. prove	6. crush
7. are	7. price	7. not	7. bad
8. there	8. box	8. frog	8. end
9. such	9. pants	9. there	9. dish
10. wheat	10. net	10. net	10. pest
11. hurt	11. then	11. soap	11. rag
12. plans	12. gift	12. ford	12. heat
13. rat	13. no	13. hid	13. are
14. dish	14. pile	14. price	14. no
15. drive	15. farm	15. ride	15. bead
16. ride	16. prove	16. pan	16. is
17. trade	17. pan	17. crash	17. gift
18. guess	18. toe	18. wheat	18. trade
19. rise	19. bead	19. farm	19. rub
20. hose	20. not	20. rise	20. hurt
21. pest	21. slip	21. eyes	21. then
22. this	22. eyes	22. plans	22. guess
23. hid	23. tone	23. bite	23. box
24. heat	24. fork	24. fork	24. range
25. frog	25. is	25. drive	25. hose
26. is	26. frog	26. hose	26. drive
27. fork	27. heat	27. range	27. fork
28. tone	28. hid	28. box	28. bite
29. eyes	29. this	29. guess	29. plans
30. slip	30. pest	30. then	30. eyes
31. not	31. hose	31. hurt	31. rise
32. bead	32. rise	32. rub	32. farm
33. toe	33. guess	33. trade	33. wheat
34. pan	34. trade	34. gift	34. crash
35. prove	35. ride	35. is	35. pan
36. farm	36. drive	36. bead	36. ride
37. pile	37. dish	37. no	37. price
38. no	38. rat	38. are	38. hid
39. gift	39. plans	39. heat	39. ford
40. then	40. hurt	40. rag	40. soap
41. net	41. wheat	41. pest	41. net
42. pants	42. such	42. dish	42. there
43. box	43. there	43. end	43. frog
44. price	44. are	44. bad	44. not
45. ford	45. crash	45. crush	45. prove
46. bite	46. crush	46. pants	46. this
47. end	47. bar	47. pile	47. rat
48. rub	48. rag	48. slip	48. such
49. smile	49. bad	49. smile	49. toe
50. soap	50. range	50. tone	50. bar

PBF-2 (LIST B)

No. 1	No. 2	No. 3	No. 4
1. grease	1. niece	1. rib	1. night
2. bounce	2. howl	2. our	2. them
3. tongue	3. flood	3. else	3. pump
4. throb	4. vest	4. tan	4. glass
5. glass	5. tank	5. throb	5. flood
6. fuse	6. boot	6. job	6. five
7. charge	7. arm	7. need	7. bud
8. hit	8. them	8. bill	8. bounce
9. bought	9. ways	9. rap	9. boot
10. wish	10. trash	10. shake	10. tank
11. bean	11. mad	11. charge	11. wish
12. cloud	12. bill	12. mad	12. arm
13. need	13. rap	13. nut	13. log
14. smash	14. else	14. smash	14. set
15. course	15. nut	15. grease	15. awe
16. quart	16. rib	16. niece	16. bought
17. loose	17. our	17. that	17. howl
18. hop	18. start	18. shoe	18. trash
19. five	19. did	19. vest	19. hit
20. tip	20. set	20. cloud	20. did
21. log	21. shoe	21. tongue	21. course
22. that	22. awe	22. tip	22. fuse
23. tan	23. job	23. start	23. ways
24. bud	24. pump	24. quart	24. loose
25. shake	25. night	25. hop	25. bean
26. night	26. shake	26. bean	26. hop
27. pump	27. bud	27. loose	27. quart
28. job	28. tan	28. ways	28. start
29. awe	29. that	29. fuse	29. tip
30. shoe	30. log	30. course	30. tongue
31. set	31. tip	31. did	31. cloud
32. did	32. five	32. hit	32. vest
33. start	33. hop	33. trash	33. shoe
34. our	34. loose	34. howl	34. that
35. rib	35. quart	35. bought	35. niece
36. nut	36. course	36. awe	36. grease
37. else	37. smash	37. set	37. smash
38. rap	38. need	38. log	38. nut
39. bill	39. cloud	39. arm	39. mad
40. mad	40. bean	40. wish	40. charge
41. trash	41. wish	41. tank	41. shake
42. ways	42. bought	42. boot	42. rap
43. them	43. hit	43. bounce	43. bill
44. arm	44. charge	44. bud	44. need
45. boot	45. fuse	45. five	45. job
46. tank	46. glass	46. flood	46. throb
47. vest	47. throb	47. glass	47. tan
48. flood	48. tongue	48. pump	48. else
49. howl	49. bounce	49. them	49. our
50. niece	50. grease	50. night	50. rib

PBF-3 (LIST C)

No. 1	No. 2	No. 3	No. 4
1. purse	1. tenth	1. cake	1. whirl
2. barge	2. take	2. wedge	2. take
3. praise	3. size	3. thick	3. stood
4. cake	4. drop	4. sad	4. oak
5. dig	5. class	5. vow	5. nest
6. deck	6. far	6. please	6. dig
7. air	7. path	7. may	7. barge
8. trip	8. clown	8. fog	8. drop
9. thick	9. nest	9. trip	9. dim
10. why	10. leave	10. leave	10. noise
11. laugh	11. town	11. tenth	11. praise
12. past	12. fresh	12. far	12. horse
13. sad	13. look	13. jam	13. air
14. dim	14. oak	14. path	14. look
15. dog	15. fame	15. shout	15. check
16. shout	16. please	16. neck	16. law
17. who	17. neck	17. deck	17. fresh
18. guard	18. wedge	18. why	18. who
19. sled	19. check	19. fame	19. size
20. late	20. may	20. sled	20. laugh
21. noise	21. stood	21. eight	21. town
22. vow	22. eight	22. past	22. guard
23. jam	23. whirl	23. class	23. clown
24. horse	24. flash	24. flash	24. purse
25. fog	25. law	25. dog	25. late
26. law	26. fog	26. late	26. dog
27. flash	27. horse	27. purse	27. flash
28. whirl	28. jam	28. clown	28. class
29. eight	29. vow	29. guard	29. past
30. stood	30. noise	30. town	30. eight
31. may	31. late	31. laugh	31. sled
32. check	32. sled	32. size	32. fame
33. wedge	33. guard	33. who	33. why
34. neck	34. who	34. fresh	34. deck
35. please	35. shout	35. law	35. neck
36. fame	36. dog	36. check	36. shout
37. oak	37. dim	37. look	37. path
38. look	38. sad	38. air	38. jam
39. fresh	39. past	39. horse	39. far
40. town	40. laugh	40. praise	40. tenth
41. leave	41. why	41. noise	41. leave
42. nest	42. thick	42. dim	42. trip
43. clown	43. trip	43. drop	43. fog
44. path	44. air	44. barge	44. may
45. far	45. deck	45. dig	45. please
46. class	46. dig	46. nest	46. vow
47. drop	47. cake	47. oak	47. sad
48. size	48. praise	48. stood	48. thick
49. take	49. barge	49. take	49. wedge
50. tenth	50. purse	50. whirl	50. cake

PBF-4 (LIST D)

No. 1	No. 2	No. 3	No. 4
1. frown	1. move	1. pipe	1. new
2. bus	2. hook	2. on	2. stitch
3. test	3. fed	3. earn	3. owl
4. stuck	4. tick	4. slap	4. food
5. food	5. sour	5. stuck	5. fed
6. float	6. bike	6. hot	6. fan
7. cart	7. aim	7. mend	7. bush
8. gas	8. stitch	8. bee	8. bus
9. blonde	9. touch	9. pinch	9. bike
10. yes	10. thin	10. raise	10. sour
11. beast	11. meat	11. cart	11. yes
12. cod	12. bee	12. meat	12. aim
13. mend	13. pinch	13. oils	13. judge
14. roe	14. earn	14. roe	14. race
15. creak	15. oils	15. gas	15. badge
16. pack	16. pipe	16. move	16. blonde
17. mat	17. on	17. stars	17. hook
18. hog	18. skid	18. rip	18. thin
19. fan	19. curse	19. tick	19. gas
20. tent	20. race	20. cod	20. curse
21. judge	21. rip	21. test	21. creak
22. stars	22. badge	22. tent	22. float
23. slap	23. hot	23. skid	23. touch
24. bush	24. owl	24. pack	24. mat
25. raise	25. new	25. hog	25. beast
26. new	26. raise	26. beast	26. hog
27. owl	27. bush	27. mat	27. pack
28. hot	28. slap	28. touch	28. skid
29. badge	29. stars	29. float	29. tent
30. rip	30. judge	30. creak	30. test
31. race	31. tent	31. curse	31. cod
32. curse	32. fan	32. gas	32. tick
33. skid	33. hog	33. thin	33. rip
34. on	34. mat	34. hook	34. stars
35. pipe	35. pack	35. blonde	35. move
36. oils	36. creak	36. badge	36. frown
37. earn	37. roe	37. race	37. roe
38. pinch	38. mend	38. judge	38. oils
39. bee	39. cod	39. aim	39. meat
40. meat	40. beast	40. yes	40. cart
41. thin	41. yes	41. sour	41. raise
42. touch	42. blonde	42. bike	42. pinch
43. stitch	43. gas	43. bus	43. bee
44. aim	44. cart	44. bush	44. mend
45. bike	45. float	45. fan	45. hot
46. sour	46. food	46. fed	46. stuck
47. tick	47. stuck	47. food	47. slap
48. fed	48. test	48. owl	48. earn
49. hook	49. bus	49. stitch	49. on
50. move	50. frown	50. new	50. pipe

BIBLIOGRAPHY

BIBLIOGRAPHY

Books

Cochran, W. G., and Cox, G. M., Experimental Designs. New York: John Wiley and Sons, 1950.

Guilford, J. P., Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Co., 1950.

_____. Psychometric Methods. New York: McGraw-Hill Book Co., 1936.

McNemar, Quinn, Psychological Statistics. New York: John Wiley and Sons, 1949.

Morley, D. E., An Analysis by Means of the Sound Spectrograph of Intelligibility Variations of Consonant Sounds Spoken by Deaf Persons. Ann Arbor: University Microfilms, 1949.

Van Riper, Charles, Speech Correction Principles and Methods. New York: Prentice-Hall, 1947.

Articles

Bangs, J. L., "A Clinical Analysis of the Articulatory Defects of the Feebleminded," Journal of Speech Disorders, VII (1942).

Black, J. W., "The Origin and Nature of the Studies," (Studies in Speech Intelligibility), Speech Monographs, XIII (1946).

Blattner, "An Experimental Study of the Testing of Pronunciation," Speech Monographs, XV (1948).

Brackett, I. P., "Intelligibility Related to Pitch," Speech Monographs, XIII (1946).

Curry, R., Kennedy, L., Wagner, L., and Wilke, W., "A Phonographic Scale for the Measurement of Defective Articulation," Journal of Speech and Hearing Disorders, VIII (1943).

Curtis, J. F., "Intelligibility Related to Microphone Position," Speech Monographs, XIII (1946).

Draegert, G. L., "Intelligibility Related to Articulation," Speech Monographs, XIII (1946).

Eagan, J. P., "Articulation Testing Methods," Laryngoscope, LVIII (September, 1948).

_____, "Transmission and Reception of Sounds Under Combat Conditions," Summary Technical Report of Division 17, OSRD III. Washington: Government Printing Office, 1946.

Falconer, G. A., and Davis, H., "The Intelligibility of Connected Discourse as a Test for the Threshold for Speech," Laryngoscope, LVII (September, 1947).

Fay, P. J., and Middleton, W. C., "Rating a Speaker's Voice When Heard Over a Public Address System," Quarterly Journal of Speech, XXXVI (December, 1951).

Fletcher, Harvey, and Steinberg, J. C., "Articulation Testing Methods," Bell System Technical Journal, VIII (1929).

French, N. R., and Steinberg, J. C., "Factors Governing the Intelligibility of Speech Sounds," Journal of the Acoustical Society of America, XIX (1947).

Haagen, C. H., "Intelligibility Measurement," Speech Monographs, XIII (1946).

Henderson, F. M., "Accuracy in Testing the Articulation of Speech Sounds," Journal of Educational Research, XXXI (1938).

Hudgins, C. V., " A Method of Appraising the Speech of the Deaf," Volta Review, (December, 1949).

_____. Hawkins, J. E., Karlin, J. E., Stevens, S. S., "The Development of Recorded Auditory Tests for Measuring Hearing Loss for Speech," Laryngoscope, LVII (January, 1947).

_____. and Numbers, F. C., "An Investigation of the Intelligibility of the Speech of the Deaf," Genetic Psychology Monographs, XXV (1942).

Latham, A. J., "Audibility and Distinctness in the Speech of High School Students," Quarterly Journal of Speech, XVII (February, 1931).

Lightfoot, C. M., "Some Effects of the Common Cold on Speech," Archives of Otolaryngology, LI (April, 1950).

Mason, H. M., "The Effect of Very Loud Speech Signals Upon Intelligibility," Speech Monographs, XIII (1946).

Moser, H, M., "Intelligibility Related to Loudness," Speech Monographs, XIII (1946).

Moore, Paul, "Intelligibility Related to Pitch," Speech Monographs, XIII (1946).

Spriestersbach, D. C., and Curtis, J. F., "Misarticulation and Discrimination of Speech Sounds," Quarterly Journal of Speech, XXXVII (December, 1951).

Steer, Mack and Hadley, J. M., "The Speech Intelligibility Program in Naval Aviation," Quarterly Journal of S Speech, XXXII (April, 1946).

Stevens, W. E., "A Study of Phonetic Factors in Relation to Accuracy of Transmission of Words in Airplane Noise," Speech Monographs, XIII (1946).

Symonds, P. M., "On the Loss of Reliability in Ratings Due to the Coarseness of the Scale," Journal of Experimental Psychology, VII (1924).

Whan, F. L., "Training in Listening and in Voice and Diction for the Airplane Pilot," Quarterly Journal of Speech, XXX (October, 1944).

Wilke, W. H., "Development and Application of a Scale for Measuring Diction," Quarterly Journal of Speech, XXIV (1938).

Williams, H. M., "A Qualitative Analysis of the Erroneous Speech Sound Substitutions of Pre-school Children," University of Iowa Studies in Child Welfare, XIII (1937).

Wood, K. S., "Parental Maladjustment and Functional Articulatory Defects in Children," Journal of Speech Disorders, XI (1946).

Unpublished Material

Haskins, Harriet, "A Phonetically Balanced Test of Speech Discrimination for Children," Unpublished Master's Thesis, Northwestern University, 1949.

Quick, Marion, "The Development of a Test of Speech Perception to Measure the Acoustic Achievement of Young Deaf Children," Unpublished Master's Thesis, Smith College, 1949.

RULES COVERING USE OF MANUSCRIPT THESES
IN THE UNIVERSITY OF MICHIGAN LIBRARY
AND THE GRADUATE SCHOOL OFFICE

Unpublished theses submitted for the doctor's degrees and deposited in the University of Michigan Library and in the Office of the Graduate School are open for inspection, but are to be used only with due regard to the rights of the authors. For this reason it is necessary to require that a manuscript thesis be read within the Library or the Office of the Graduate School. If the thesis is borrowed by another Library, the same rules should be observed by it. Bibliographical references may be noted, but passages may be copied only with the permission of the authors, and proper credit must be given in subsequent written or published work. Extensive copying or publication of the thesis in whole or in part must have the consent of the author as well as of the Dean of the Graduate School.

This thesis by .....
has been used by the following persons, whose signatures attest their acceptance of the above restrictions.

A Library which borrows this thesis for use by its readers is expected to secure the signature of each user.

NAME AND ADDRESS	DATE

Form 6327 12-50 1M