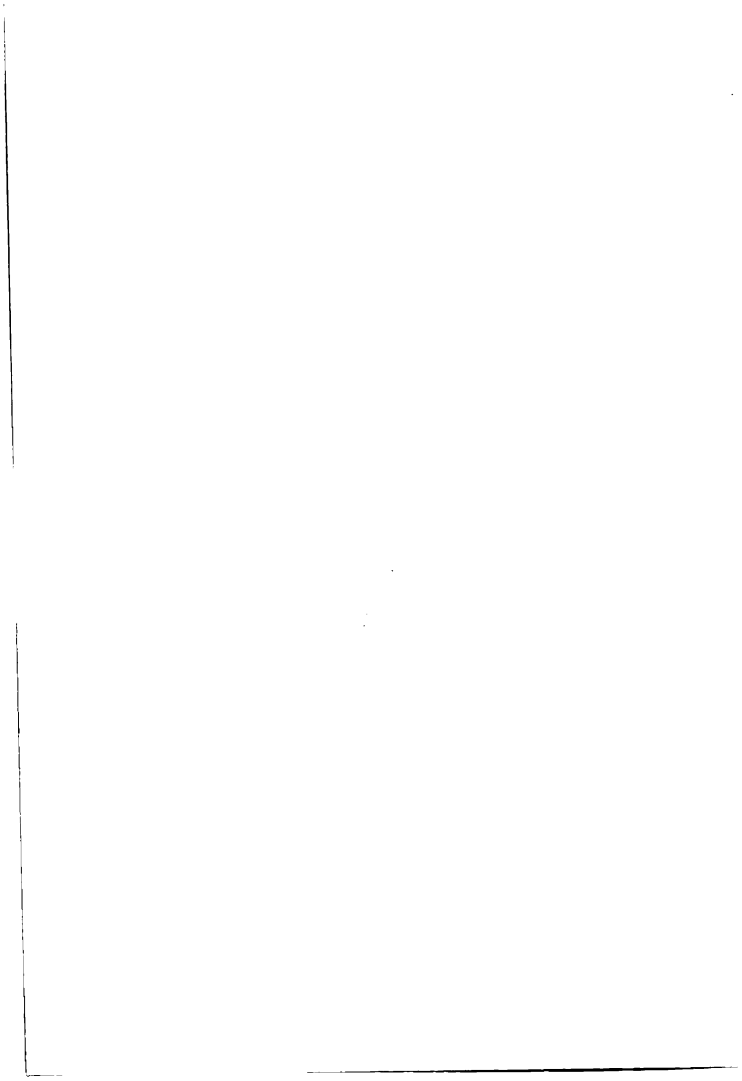
A COMPARISON OF C. I. D. AUDITORY TEST W-22 AND C. I. D. EVERYDAY SPEECH

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY

Joan B. Dalio

1965



ABSTRACT

A COMPARISON OF
C.I.D. AUDITORY TEST W-22
AND
C.I.D. EVERYDAY SPEECH

by Joan B. Dolio

Determination of how an individual's hearing is functioning in everyday conversational situations plays an important part in current audiometric hearing evaluations. At the present time, word lists are used to determine the ability to hear and understand speech. Since we do not usually speak in single words, it was thought that a measure more closely approximating everyday conversation would be a more reliable means of testing the function of hearing in social situations. One way to do this would be to use a test which utilizes sentences as test material.

C.I.D. Everyday Speech is a group of 100 sentences divided into lists of ten sentences each. These sentences were constructed to be representative of everyday speech. A recording of these sentence lists was made using five male and five female speakers in the age range from 20 to 70 years.

The purpose of this study was to compare these sentence lists with the currently used C.I.D. PB word lists. The following null hypotheses were tested: (1) There is

no correlation between scores obtained on C.I.D. Auditory Test W-22 and scores obtained on C.I.D. Everyday Speech lists; (2) The reliability coefficient for C.I.D. Auditory Test W-22 as used in this study is less than .95; (3) The reliability coefficient for the C.I.D. Everyday Speech lists is less than .95; (4) The reliability for the C.I.D. Everyday Speech lists is less than or equal to the reliability of C.I.D. Auditory Test W-22.

Forty normal-hearing subjects between the ages of 20 and 40 years were tested in a sound isolated chamber at the Michigan State University Speech and Hearing Clinic. One form of C.I.D. Auditory Test W-22 and one list of C.I.D. Everyday Speech was presented to each subject via an Allison Model 22 Clinical and Research Audiometer. A signal to noise ratio of -23 dB was used in order to depress scores. The signal was presented at 60 dB SPL, and the white noise at 83 dB SPL.

A product-moment correlation coefficient was computed to determine to what extent the sentence lists are related to the word lists. A positive correlation of .43 was found. Since the r of .43 was significantly greater than zero, as demonstrated by the .05 confidence interval, the null hypothesis of no correlation was rejected.

Odd-even reliability coefficients were computed by the product-moment method for both tests and were corrected for length using the Spearman-Brown formula. A

reliability coefficient of .95 was obtained for C.I.D. Auditory Test W-22. For C.I.D. Everyday Speech lists, a reliability coefficient of .98 was found. Since the confidence interval at the .05 level for the r of C.I.D. Auditory Test W-22 had as its lower limit an r of .92, the null hypothesis could not be rejected. The reliability coefficient computed for C.I.D. Everyday Speech was .98. The lower limit of this confidence interval, an r of .97, allowed rejection of the null hypothesis. Although both reliability coefficients were extremely high, the sentence lists were more reliable than the word lists as indicated by the .05 confidence intervals of their respective reliability coefficients. Because of this, the null hypothesis stating that the reliability coefficient for C.I.D. Everyday Speech lists is less than or equal to the reliability of C.I.D. Auditory Test W-22 was rejected.

Because of the low positive correlation found between the two tests as a whole, it was concluded that these two tests measure only somewhat similar--certainly not the same--kind of behavior. This would seem to supply a reason for the utilization of the sentence lists in audiometric hearing evaluation.

A COMPARISON OF

C.I.D. AUDITORY TEST W-22

AND

C.I.D. EVERYDAY SPEECH

By

Joan B. Dolio

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Speech

1965

TABLE OF CONTENTS

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

> The most important function of hearing for modern man is in auditory communication. What the deaf man wants most is to hear human speech, so that he may understand. The other functions, such as warning of danger, enjoyment of music, controlling the quality and loudness of his own speech, and locating and identifying miscellaneous sounds and signals, are secondary to the hearing of speech or closely related to it.[1]

This desire of modern man to hear and understand the speech of others is taken into consideration in audiometric hearing and hearing aid evaluations.

At the present time an audiometric or hearing-aid evaluation consists of two parts: pure-tone audiometry and speech audiometry. Through pure-tone audiometry the functioning of the auditory mechanism can be defined.

> One of the reasons for considering the measurement of a person's ability to hear speech is to fill in the gap between the audiogram and the person's ability to communicate with his fellows in everyday life.[2]

With speech audiometry the audiologist can gain additional

---

[1]Hallowell Davis, "The Articulation Area and the Social Adequacy Index for Hearing," Laryngoscope, VIII (1948), p. 761.

[2]Leland A. Watson and Thomas Tolan, Hearing Tests and Hearing Instruments (Baltimore: The Williams & Wilkins Company, 1949), p. 64.

information concerning the client's hearing handicap. It is generally agreed that the pure tone test results plus speech test results give a more reliable and complete picture of hearing function than either of them does by itself.

In aural rehabilitation the most important factor appears to be how well speech is heard and understood by the client. Speech audiometry allows the audiologist to obtain information concerning the client's understanding of speech and assists him in determining how well the client will function in social situations--situations he will encounter in daily living.

In order to determine how the client is hearing speech, Hirsh suggests that the audiologist answer the following questions: "(1) How intense must a particular sample of speech be in order that it be just intelligible to a listener? (2) At a given intensity, how intelligible is a given sample of speech?"[3] The first question deals with the measurement of a threshold of intelligibility for speech and considers the intensity as a variable. This measure has been named Speech Reception Threshold. The second question is concerned with the intelligibility of a sample of speech at one fixed intensity. This measure is called discrimination. If both of the questions are adequately answered, the audiologist has obtained information

---

[3]Ira J. Hirsh, The Measurement of Hearing (New York: McGraw-Hill Book Company, Inc., 1952), p. 127.

concerning how well speech is heard and understood by his client.

In this study the area of concern is discrimination testing. With many clients, the primary problem is a reduction of speech intelligibility--the ability to discriminate. Speech discrimination is, audiologically, sometimes called articulation. An articulation test, therefore, is a test which determines the ability of a client to discriminate among given samples of speech.

The current method of measuring discrimination function is through administration of phonetically balanced (PB) word lists at levels considerably above the client's Speech Reception Threshold. It is, therefore, a supra-threshold measure. Each PB word list is comprised of 50 monosyllabic words which contain samples of speech sounds in roughly the same proportion that they occur in English speech. These word lists are adapted from the Harvard Psycho-Acoustics Laboratory PB word lists and meet certain criteria of familiarity. PB word lists are available in recorded form. One of these tests was prepared at Central Institute for the Deaf in Saint Louis, Missouri, and is called C.I.D. Auditory Test W-22. In all the PB word lists there are 200 words which have been arranged into 24 different orders. The 200 words, as they appear on C.I.D. Auditory Test W-22, Lists 1A, 2A, 3A, and 4A, are listed in Appendix C. Because these lists are phonetically balanced

with the occurrence of each sound in English speech and with each other, Newby states that there should be very little variation in discrimination score from list to list.[4]

Since C.I.D. Auditory Test W-22 is currently the most commonly used clinical measure of discrimination function, this researcher has chosen to compare results obtained from its use with results obtained from another measure of discrimination function. This second measure is C.I.D. Everyday Speech, which is a list of 100 sentences divided into 10 different groups. These sentences are constructed so as to be representative of American speech. The criteria for the development of a measure of 'everyday speech' were developed by a working group of the Armed Forces-National Research Council Committee on Hearing and Bio-Acoustics (CHABA). Because C.I.D. Everyday Speech meets the requirements outlined by CHABA,[5] it should give a good indication of a client's hearing in everyday life.

Statement of the Problem

The purpose of this study is to compare C.I.D. Auditory Test W-22 and C.I.D. Everyday Speech to determine the extent to which these two tests of discrimination

---

[4]Hayes A. Newby, Audiology (New York: Appleton-Century-Crofts, 1964), p. 115.

[5]Richard Silverman and Ira J. Hirsh, "Problems Related to the Use of Speech in Clinical Audiometry," Annals of Otology, Rhinology and Laryngology, LXIV (1955), p. 1243.

correlate.

The following questions and null hypotheses have been formulated to be explored in this study:

(1) Are these two tests of discrimination measuring similar behavior? If they are, the scores obtained should correlate to some extent.

Null hypothesis: There is no correlation between scores obtained on C.I.D. Auditory Test W-22 and scores obtained on C.I.D. Everyday Speech lists.

(2) Is the reliability of C.I.D. Auditory Test W-22, Lists 1A, 2A, 3A, and 4A high enough to yield consistent test results? According to Greene, a correlation coefficient of .95 or greater is the criterion for a "reliable test."[6]

Null hypothesis: The reliability coefficient for C.I.D. Auditory Test W-22 as used in this study is less than .95.

(3) Is the reliability of the C.I.D. Everyday Speech lists high enough to meet Greene's criterion for a "reliable test"?

Null hypothesis: The reliability coefficient for the C.I.D. Everyday Speech lists is less than .95.

(4) Is the reliability of C.I.D. Auditory Test W-22 higher than the reliability of the C.I.D. Everyday Speech lists?

Null hypothesis: The reliability for the C.I.D. Everyday

---

[6]Edward B. Greene, Measurement of Human Behavior (New York: The Odyssey Press, 1952), p. 49.

Speech lists is less than or equal to the reliability of C.I.D. Auditory Test W-22.

Definition of Terms

In this study the terms used are defined as follows:

1. Discrimination Test--An evaluation of a subject's ability to identify and repeat certain samples of speech.
2. Phonetically Balanced (PB) Word Lists--Lists of 50 monosyllabic words chosen so that each list will include samples of speech sounds in the same proportion in which they occur in the English language.[7]
3. C.I.D. Auditory Test W-22--Recorded forms of the phonetically balanced word lists. Each list consists of 50 monosyllabic words arranged in six different word orders.
4. C.I.D. Everyday Speech--A group of 100 sentences divided into 10 lists. These sentences were constructed at Central Institute for the Deaf, St. Louis, Missouri, and are representative of everyday speech. Scoring of these sentences is based on 50 key words in each list of 10 sentences.
5. Word Discrimination Score--The percentage of words correctly identified by a subject when C.I.D. Auditory Test W-22 is presented.
6. Sentence Discrimination Score--The percentage of key

---

[7]Newby, op. cit., p. 115.

words correctly identified by a subject when C.I.D. Everyday Speech lists are presented.

7. Spondaic Words--Two syllable words pronounced with equal stress on each syllable.
8. Speech Reception Threshold--The intensity level at which a subject repeats correctly 50 per cent of the spondaic words presented.
9. Normal-Hearing Subjects--Subjects who claim to have no hearing loss.
10. Intensity Level--All intensity levels presented in this study are reported in terms of sound pressure level, unless otherwise stated.

Importance of the Study

One of the reasons for using PB word lists to measure intelligibility is that they are smaller units of speech than sentences or everyday conversation. Because of this, "more precise relations between intelligibility and the physical dimensions of the test unit"[8] can be determined. Intensity is not as easily controlled when using samples of everyday speech if naturalness of delivery remains as it is using smaller units of speech. It would seem that in spite of this a more valid measure for determining discrimination function and the ability to function in everyday situations would be a discrimination measure consisting

---

[8]Hirsh, op. cit., p. 128.

of common everyday sentences. In regard to this Hirsh has stated:

> To be sure, the relation between such lists [the PB word lists] and the continuous flow of words we encounter in conversation is not very clear. Instead, therefore, we may attempt to devise a more valid test by using groups of words that might appear in conversation. One such group is the sentence.[9]

Many groups of sentences have been prepared. One such collection of lists, C.I.D. Everyday Speech, was developed at Central Institute for the Deaf, St. Louis, Missouri. These are the sentence lists to be used in this study.

This study was undertaken because of the problems involved in using the 50-word PB lists. Dissatisfaction has recently been expressed in regard to these lists including the length of time necessary for administration, the disagreement about the level at which to administer them, the difficulties involved in hearing aid evaluation when several lists must be administered, and questions regarding the validity and reliability of the word lists.

Organization of the Study

This thesis is organized in five chapters. Chapter I includes a statement of the problem, background information, and definition of the terms for this study. Chapter II contains a summary of the literature in the field of speech audiometry. Procedures, equipment, and subjects

---

[9] Ibid., p. 131.

will be discussed in Chapter III. Chapter IV contains the results of the study, and Chapter V is a brief discussion of conclusions and implications for future research in this particular area.

CHAPTER II

REVIEW OF THE LITERATURE

There are many tests of hearing for speech. Some of them are crude and not standardized. These tests, such as the conversational voice test, when used by experienced clinicians, can be of great practical value. Since tests such as these are not ordinarily used in hearing evaluations in speech and hearing clinics throughout this country, they will not be included in the review of the literature. The tests and test materials discussed in this review of the literature will be those tests of hearing level for speech and auditory discrimination ability that are currently being used in hearing evaluations.

The speech tests used at this time in speech and hearing clinics include the following:

1. Speech Reception Threshold (SRT)--The hearing level at which a client can repeat 50 per cent of the spondaic words presented correctly. Zero SRT was established by ASA (American Standards Association) at a sound pressure level of 22 dB.[1]

---

[1]American Standards Association, Specifications for Speech Audiometers, Z24.13-1953 (New York: American Standards Association, Inc.).

2. Most Comfortable Loudness (MCL)--The hearing level at which speech is most comfortable for the client as measured by means of running speech.[2]

3. Detection Threshold--The level at which the client is able to detect the presence of sound 50 per cent of the time.

4. Tolerance Level--The hearing level at which speech becomes uncomfortably loud. "The purpose of this measure is to find the upper limit of the patient's range of hearing for speech. . . ."[3]

5. Maximum Discrimination Score (PB Max)--This test usually utilizes phonetically balanced word lists "presented at a level which will be at least 40 db above the patient's previously determined SRT."[4] The correct responses are usually stated as a percentage.

Probably the two most important of these tests are Speech Reception Threshold which utilizes the Spondee Word Lists and Maximum Discrimination Score which uses phonetically balanced word lists.

The Development of Speech Audiometry

The sounds of speech have come to occupy an important place among the auditory stimuli that are used in

---

[2]Newby, op. cit., p. 113.

[3]Ibid., p. 114.

[4]Ibid., p. 115.

clinical audiometry. According to Davis and Silverman, the Western Electric 4C group audiometer established the principle of speech audiometry. This audiometer was used originally for school screening tests and later was used for screening tests during World War II. It provided a single measure of hearing, was reliable, and was efficient in regard to time necessary for administration. Words or sentences were used as the test materials.[5]

Some of the first tests of speech were developed at the Bell Telephone Laboratories for the purpose of testing telephone equipment. Two of the early articulation tests involved the repetition of numbers and the repetition of nonsense syllables. From the original nonsense syllable test Fletcher and Steinberg developed an articulation test by modification of the International Phonetic Association Alphabet. The alphabet as they revised it consisted of 48 simple sounds: 24 consonants, 19 vowels, and five diphthongs.[6] They combined the fundamental sounds into syllables in a variety of ways but adhered to very simple syllable forms. Each test consisted of a list of 66 syllables.[7] They also developed lists of common English words

---

[5]Hallowell Davis and S. Richard Silverman, Hearing and Deafness (New York: Holt, Rinehart, and Winston, Inc., 1961), p. 181.

[6]H. Fletcher and J. C. Steinberg, "Articulation Testing Methods," Bell System Technical Journal, VIII (1929), p. 810.

[7]Ibid., p. 811.

arranged in the consonant-vowel-consonant sequence. Two lists were used, one for the evaluation of vowel intelligibility and the other for consonant intelligibility. On the first list only the vowel errors were counted, and on the second only the consonant errors. Sentence lists were also developed at the Bell Telephone Laboratories during the same period.

> During World War II, considerable effort was expended in the development of articulation testing methods for the evaluation of various types of military communications equipment. It turned out that certain of these tests, developed at the Psycho-Acoustics Laboratory, Harvard University, were applicable to the clinical evaluation of hearing.[8]

In order to develop tests more appropriate to audiometric testing, a program was undertaken at the Harvard Psycho-Acoustics Laboratory with the following aims:

> 1. To explore further the problems involved in the construction of audiometric tests for measuring directly the hearing loss for speech.
>
> 2. To produce a test suitable for precise laboratory measurements of all degrees of hearing loss.
>
> 3. To explore by means of verbal tests the possibility of differentiating between high-frequency deafness and deafness which is uniform throughout the audible frequency range.[9]

The results of this program were PAL Auditory Test No. 9

---

[8]Hirsh et al., "Development of Materials for Speech Audiometry," Journal of Speech and Hearing Disorders, XVII (1952), p. 321.

[9]Hudgins et al., "The Development of Recorded Auditory Tests for Measuring Hearing Loss for Speech," Laryngoscope, LVII (1947), p. 62.

and PAL Auditory Test No. 12. PAL Auditory Test No. 9 was Hearing Threshold for Words; and PAL Auditory Test No. 12, Threshold of Hearing for Sentences.

PAL Auditory Test No. 9 consists of two lists of 42 dissyllabic words of the spondee stress pattern. "The high audibility of the spondee as compared to other dissyllabic words . . . is due apparently to the difference in stress patterns."[10] Both syllables are very nearly equally stressed.

The second PAL test consisted of eight lists of short simple questions. "The lists are composed of 28 items divided into seven groups of four items each,"[11] and each group is recorded 4 dB less intense than the preceding group.

The criteria for this program consisted of the following: (1) familiarity; (2) phonetic dissimilarity; (3) normal sampling of English speech sounds; and (4) homogeneity with respect to basic audibility.

Also at the Psycho-Acoustics Laboratory at Harvard University Egan developed the PB-50 word lists. The vocabulary for this test was drawn from 1200 monosyllabic words. The exceptionally easy words were discarded. Twenty-four lists of 50 words were constructed. "The words were assigned to each list on the basis of the composition of the

---

[10] Ibid., p. 65.

[11] Ibid., p. 67.

first part of the word."[12] Care was taken so that these lists would be more phonetically balanced than previous word lists. The following were the criteria for the development of these lists: (1) monosyllabic structure, (2) equal average difficulty, (3) equal range of difficulty, (4) equal phonetic composition, (5) a composition representative of English speech, and (6) words in common usage.

> Resemblance to the English language is not entirely intact as a result of compromises and the resemblance of the phonetic structure of the lists to that of the language is not entirely intact.[13]

C.I.D. Auditory Tests W-1 and W-22

A revision of the Harvard tests was undertaken at Central Institute for the Deaf, St. Louis, Missouri, because it was felt that there were deficiencies with respect to clinical use in these tests. Certain recordings of PAL Auditory Test No. 9 yielded slightly different thresholds than other recordings. The 1200 word PB-50 vocabulary list was felt to be too large. The PB lists were not available in standardized recorded form.

Two basic improvements were made at Central Institute for the Deaf. The vocabulary for both the spondee lists and the PB lists was restricted to include only words which met certain criteria of familiarity. The PB

---

[12] James P. Egan, "Articulation Testing Methods," Laryngoscope, LVIII (1948), p. 963.

[13] Ibid., p. 964.

vocabulary was more rigidly phonetically balanced. The second improvement made was a recording on magnetic tape.

The improved version of PAL Auditory Test No. 9 is C.I.D. Auditory Test W-2 which is spondees recorded at descending levels of intensity. PAL Auditory Test No. 14 was revised, and it became C.I.D. Auditory Test W-1 which consists of spondees recorded at a constant intensity level. The modified PB-50 lists became C.I.D. Auditory Test W-22.

C.I.D. Auditory Test W-1

This test consists of six scramblings of 36 spondaic words recorded at a constant intensity level. Each word is presented at an intensity level 10 dB lower than the level of the carrier phrase.

> The Articulation score rises from 0 to 100% within a range of about 20 db. There is an increase from 20 to 80% within a range of 8 db and throughout this range the slope or rate of rise in score is about 8% per db. Since the threshold falls on the steepest part of the function, it is crossed very abruptly and, therefore, it can be very sensitively determined with this test.[14]

In the original standardization for both experienced and inexperienced listeners the absolute thresholds were 20 dB and 21 dB re 0.0002 microbar.

C.I.D. Auditory Test W-2 consists of the same spondee words. The intensity is attenuated 3 dB every three words.

---

[14]Hirsh et al., op. cit., pp. 325-26.

C.I.D. Auditory Test W-22

The vocabulary for this test consists of 200 monosyllabic words which are divided into four different word lists containing 50 words each. Six scramblings of each list have been recorded. Each list is phonetically balanced. The test is used to determine a client's discrimination score or discrimination loss for speech. The criteria for the development of these lists were: all the words had to be one syllable with no repetition of words in the different lists; any word used had to be a familiar word; and the phonetic composition of each word list had to correspond to English as a whole as closely as possible.[15]

The original data in this study showed no consistent differences between discrimination scores on the four different lists.

In comparison to the Harvard PB-50 word lists Hirsh stated the intelligibility for the older recordings (the Harvard lists) was lower at any given intensity than for the W-22 form. The words on the W-22 form were monitored so that they were much closer to each other in intensity than they were on the Harvard PB-50 lists.[16]

Phonetic Balance

Several authors have recently questioned the assumed

---

[15] Ibid., p. 328.

[16] Ibid., pp. 334-35.

importance of using phonetic balance in testing auditory discrimination. As early as 1929 Fletcher and Steinberg stated the following in regard to their word lists:

> The results obtained with these lists . . . are as representative of speech as the results that would be obtained with lists employing particular sound combinations in proportion to their frequencies of occurrence in speech.[17]

The word lists used by Fletcher and Steinberg were randomly constructed.

Elpern states,

> . . . it seems to be the consensus that phonetic balance is not as crucial a factor to the sensitivity of these tests [discrimination tests] as it was thought to be at the time the original PB-50 lists were developed.[18]

He found that half-lists were as reliable measures of discrimination as a full list of 50 words from the C.I.D. Auditory Test W-22 forms.

Campanelli divided each of the Harvard PB-50 word lists into two lists of 25 words each. The results indicated that the 25-word list gave a reliable discrimination score.

Some authors, however, have written in defense of phonetic balance. Grubb believes that phonetic balance gives face validity to the PB word lists. She states that "the PB characteristic of the whole list is lost in the

---

[17]Fletcher and Steinberg, op. cit., p. 834.

[18]B. S. Elpern, "The Relative Stability of One-Half-List and Full-List Discrimination Tests," Laryngoscope, LXXI (1961), p. 31.

half-list."[19] In a letter to the editor she states:

> The difficulty with present thinking in using one-half of a discrimination test instead of the whole list is that the whole list is being used as a sort of criterion when it is actually a fallible measure subject to the same kind of error as the half-list.[20]

She argues in defense of using whole lists even though high correlation coefficients have been obtained between whole and half lists. She does this because of the following:

> A sub-test score will usually correlate positively to some degree with the whole test score of which the sub-test is a part. The reason obviously is that the whole score includes the part score.[21]

Grubb also argues that the reliability of the half-list as a test of discrimination has not been established.

Tobias, in a letter to the editor, in answer to Grubb's criticisms of the use of half lists, states

> She wants to maintain phonetic balance despite the overwhelming clinical and experimental experience that indicates phonetic balance to be an interesting but unnecessary component of one of our current audiometric tests. . . . Of course, one must ask whether half-list tests measure the same thing as full list tests. From the literature and from reports of audiologists using half lists, one must conclude that they do. On what grounds then can one insist upon phonetic balance

---

[19] P. Grubb, "A Phonemic Analysis of Half-List Speech Discrimination Tests," Journal of Speech and Hearing Research, VI (1963), p. 271.

[20] P. Grubb, "Some Considerations in the Use of Half-List Speech Discrimination Tests," Journal of Speech and Hearing Research, VI (1963), p. 296.

[21] Ibid., p. 295.

as a criterion of discrimination test validity?[22]

Speech Intelligibility

Speech sounds are divided into two main categories: vowels and consonants. According to Hirsh, consonants are responsible for intelligibility in single English words.

Intensity also appears to affect intelligibility. In general, as intensity increases intelligibility increases.

The number of syllables in the material presented affects intelligibility. A word of two syllables would be easier to identify than a word of one syllable. A sentence would be more intelligible than a single word. When using a sentence as the material to be presented context must be considered as a variable affecting intelligibility. Because of the information gained from the context of a sentence the number of possible choices is reduced, thereby increasing the probability of correct identification of the material presented.

Distortion and background noise tend to have detrimental effects upon intelligibility. In general, as the background noise increases in intensity, intelligibility decreases. "Acoustic distortion is of three forms: frequency distortion, phase distortion, and amplitude distortion."[23]

---

[22]J. V. Tobias, "On Phonemic Analysis of Speech Discrimination Tests," Journal of Speech and Hearing Research, VII (1964), pp. 99-100.

[23]Ernest G. Wever and Merle Lawrence, Physiological Acoustics (Princeton, New Jersey: Princeton University Press, 1954), p. 118.

Frequency distortion consists of variation in sensitivity to different frequencies. In phase distortion there are different periods of delay for transmission for certain sounds or tones. Amplitude distortion refers to that which occurs due to differences in efficiency of the transmitting mechanism.[24] All of these types of distortion have effects upon intelligibility. Different types of hearing loss can affect distortion and in turn can affect intelligibility.

Several studies have been conducted in the area of speech intelligibility. Miller, Heise and Lichten report the variables of talkers, listeners, test materials, and communication equipment as affecting intelligibility. They state that one of the important variables is the range of alternatives from which a response must be selected. The threshold of intelligibility for sentences is 6 dB lower than that of isolated words.[25] They stated that this is because the range of responses is restricted.

Falconer and Davis established listeners' speech reception thresholds using connected discourse. The listener attenuated the speech until he could just hear and understand the message. This threshold compared favorably with **PAL** Auditory Test No. 9. The absolute threshold mean

---

[24] Ibid.

[25] G. A. Miller, G. Heise, and W. Lichten, "The Intelligibility of Speech as a Function of the Context of the Test Materials," Journal of Experimental Psychology, XLI (1951), p. 334.

for normal listeners was 23.23 ± 3.77 dB re 0.0002 dyne/cm^2.[26]

Everyday Sentences

The sentences to be used in the present investigation were developed at Central Institute for the Deaf. Silverman and Hirsh stated a need for developing these sentences in 1955.

> . . . we face the requirement that a test shall assist us in telling how an individual hears in everyday life. It may turn out that the threshold audiogram may be the best predictor of this ability. On the other hand, it seems reasonable to suppose that since everyday hearing is characterized largely by the hearing of speech, that speech may be a more valid predicting test material. But we do not expect, necessarily, that our lists of monosyllabic words, which aid us in diagnosis would predict this ability very well, since such words can hardly be called representative of everyday speech. . . .
>
> We must now forget the concepts of hearing tests and approach the problem from the point of view of attempting to characterize samples of everyday speech.[27]

This problem was outlined and given to the Armed Forces-National Research Council Committee on Hearing and Bio-Acoustics (CHABA) for investigation. The problem was turned over to a working group to formulate a set of criteria for representing everyday speech. The assumption was that the

---

[26]G. A. Falconer and H. Davis, "The Intelligibility of Connected Discourse as a Test for the 'Threshold for Speech,'" Laryngoscope, LVII (1947), pp. 581-87.

[27]S. Richard Silverman and Ira J. Hirsh, "Problems Related to the Use of Speech in Clinical Audiometry," Annals of Otology, Rhinology, and Laryngology, LXIV (1955), p. 1241.

sample item should be the sentence. The criteria eventually established for the sentence lists are recorded in Appendix A.

Finally, a set of 100 sentences was developed at Central Institute for the Deaf. The set is divided into ten lists of ten sentences each. The sentences as listed in Davis and Silverman, Appendix 9, are shown in Appendix B of this thesis.

In 1965 Kenneth R. Johnson developed an articulation curve for his own recording of the C.I.D. Everyday Speech lists. One male and one female talker was used in each ten-year age group from 20 to 70 years. In each list one sentence was spoken by each talker. They were instructed to read the sentences with as much normal expression as possible. Each of 30 subjects listened to the 100 sentences in counter balanced order at varying intensities. The curve of the mean articulation scores was described by Johnson as follows:

> The curve is a smooth S shape with a steep slope of about seven per cent per dB at the center. This general shape is not unlike those reported in Stevens and in Hirsh for spondees, monosyllables and PB's. In fact, it corresponds very closely with the curve for the CID Auditory Test W-1 reported by Hirsh et al.[28]

The threshold for the sentence curve in this particular

---

[28]Kenneth R. Johnson, "Construction of an Articulation Curve for Recorded Sentences" (unpublished Master's thesis, Department of Speech, Michigan State University, 1965), p. 35.

study was 23.5 dB re 0.0002 microbar, which compares favorably with the threshold for spondees (20 and 21 dB re 0.0002 microbar) and with the threshold for PB's (24 dB re 0.0002 microbar).[29]

---

[29] Ibid., pp. 35-36.

CHAPTER III

PROCEDURES AND EQUIPMENT

In this study 40 subjects were tested. A description of the procedures employed, the subjects, the equipment and test materials follows.

Subjects

Forty subjects were employed in this study. The sample consisted of 18 males and 22 females between the ages of 20 and 40 years. All subjects had normal hearing. The criteria for normal hearing were that each subject stated that his hearing was within normal limits and that he never experienced difficulty in everyday conversational situations. None of the subjects was familiar with the C.I.D. Everyday Speech lists or with the testing procedures.

Equipment

All test materials and directions were delivered through an Allison Model 22 Clinical and Research Audiometer. The C.I.D. sentence lists were administered through a Viking Model 87 tape play-back at a tape speed of seven and one-half inches per second. A Bogen Model 62 phonograph was used to deliver C.I.D. Auditory Test W-22.

Materials

Each subject listened to one form of C.I.D. Auditory Test W-22 (see Appendix C) and to one list of C.I.D. Everyday Speech (see Appendix B).

C.I.D. Auditory Test W-22

Four forms of C.I.D. Auditory Test W-22 were administered. These were tests 1A, 2A, 3A, and 4A. The other 20 forms of this test are scramblings of the words presented in these four lists. Phonograph recordings made at Telephonic Studios in St. Louis, Missouri, by Central Institute for the Deaf were employed.

C.I.D. Everyday Speech

The sentences used were constructed at Central Institute for the Deaf according to criteria set up by a Working Group of the Armed Forces-National Research Council Committee on Hearing and Bio-Acoustics (see Appendix A). The 100 sentences, divided into ten lists, are constructed to be representative of everyday speech. The recording of these sentences was done by Kenneth R. Johnson at Michigan State University. Five male and five female talkers ranging in age from 24 through 68 years were utilized. One sentence in each list was spoken by each speaker. It was felt that a recording with a different talker for each sentence in a group of ten would make the sentences more representative of everyday speech. In conversational

situations in everyday life a listener would have to listen to more than one person in order to follow a conversation. Variations due to talker intelligibility are compensated for in that each talker appears on every sentence list. The original recording of these lists was done with Scotch Low Noise 202 recording tape on an Ampex PR 10 tape recorder.

Test Procedures

The forty subjects were tested individually in a sound-isolated chamber produced by the Suttle Equipment Corporation. The subjects sat in a chair facing both speakers which were located at either corner of the room. The signal was fed into Speaker B on Channel 2 via the Allison Model 22 Clinical and Research Audiometer, and white noise was fed through Speaker A via Channel 1 in the same manner.

White noise was employed in order to depress scores and give some semblance of hearing impairment. The level of white noise was determined through a pilot study involving four subjects. A level was found which masked approximately 50 per cent of the signal. A signal-to-noise ratio of -23 dB was employed in the testing. The signal was presented to all subjects at 60 dB Sound Pressure Level (SPL) and the white noise at 83 dB SPL.

The lists were presented according to the schedule shown in Appendix D.

Before testing started, each subject was asked the

following questions:

(1) Do you consider your hearing to be "normal"?

(2) Do you ever encounter any difficulty in understanding speech in everyday situations?

The first question had to be answered affirmatively, and the second negatively before the subject was included in the study.

Each subject received the following instructions before the first test (C.I.D. Auditory Test W-22) was administered:

> You are going to hear a list of 50 English words. The voice on the record will say, "You will say . . ." You are to write the word that you hear. At the same time the words are presented noise will also be presented. The noise is going to make some of the words extremely hard to hear. The next thing you will hear is "Are you ready"? and the test will begin.

Upon completion of C.I.D. Auditory Test W-22, the following instructions were given for C.I.D. Everyday Speech:

> You are about to hear ten sentences. I would like you to write each sentence or part of a sentence that you hear. When I signal you through the window you should listen because a sentence will be presented. You will have plenty of time to write the sentence. When you are finished writing be sure to look up so I can signal you when another sentence will be presented. Are there any questions?

For C.I.D. Auditory Test W-22 the VU meter was adjusted to read 0 when the calibrating tone was presented. Each list was presented at 60 dB SPL.

The VU meter was set at -5 when the calibrating tone for the sentence lists was presented.

> Because the median intensity of each list was 70 dB SPL and the calibrating tone was originally recorded at 75 dB, the playback system was now reproducing the sentences 10 dB more intense than their median intensity. Therefore, in order to play the sentences back at levels relative to their recorded intensity, 10 dB was subtracted from the attenuator reading.[1]

Each list of C.I.D. Everyday Speech was presented at 60 dB SPL.

On both tests the number of correct responses was calculated on the basis of 50 words. These were the 50 PB words presented in C.I.D. Auditory Test W-22 and the 50 key words in each sentence list. The total number of correct words on each list was multiplied by two. This gave the percentage correct for each list. The answer forms used in this study are shown in Appendix E.

The average time necessary for the completion of both tests was approximately 15 minutes per subject.

---

[1] Ibid., p. 32.

CHAPTER IV

RESULTS AND DISCUSSION

Results

Each of the 40 subjects originally obtained two scores, one for C.I.D. Auditory Test W-22 and one for C.I.D. Everyday Speech. The scores are expressed in terms of percentage of the 50 words on each list correctly identified. Each row of Appendix F represents one subject and each column represents the scores obtained on each test.

The following null hypotheses were tested:

(1) There is no correlation between scores obtained on C.I.D. Auditory Test W-22 and scores obtained on C.I.D. Everyday Speech lists.

(2) The reliability coefficient for C.I.D. Auditory Test W-22 as used in this study is less than .95.

(3) The reliability coefficient for the C.I.D. Everyday Speech lists is less than .95.

(4) The reliability for the C.I.D. Everyday Speech lists is less than or equal to the reliability of C.I.D. Auditory Test W-22.

The results were tabulated and were treated statistically by means of a product-moment correlation to determine

the relationship between C.I.D. Auditory Test W-22 and C.I.D. Everyday Speech. A product-moment correlation coefficient of .43 was found between the two tests. This would indicate a slight correlation between the two measures.

Odd-even reliability coefficients were computed by the product-moment method for both tests. The reliability coefficients were corrected for length using the Spearman-Brown formula.[1] A reliability coefficient of .95 was obtained for C.I.D. Auditory Test W-22. For C.I.D. Everyday Speech lists a reliability coefficient of .98 was found.

Fisher's z_r transformation formula[2] was used to compute confidence intervals for the correlation coefficients. The confidence intervals obtained are as follows:

(1) Between C.I.D. Auditory Test W-22 and C.I.D. Everyday Speech the confidence interval at the .05 level of confidence obtained for an r of .43 was from .14 to .65.

(2) The confidence interval at the .05 level of confidence for the reliability coefficient of .95 for C.I.D. Auditory Test W-22 was from .92 to .98.

(3) For the forms of C.I.D. Everyday Speech the confidence interval for an r of .98 at the .05 level of confidence was from .97 to .99.

---

[1]George Ferguson, Statistical Analysis in Psychology and Education (New York: McGraw-Hill Book Company, Inc., 1959), p. 280.

[2]Virginia Senders, Measurement and Statistics (New York: Oxford University Press, 1958), pp. 477-78.

Discussion

The product-moment correlation coefficient was computed to determine to what extent the sentence lists are related to the word lists. A positive correlation of .43 was found. Since the *r* of .43 is significantly greater than zero as demonstrated by the .05 confidence interval, it can be said that these two discrimination tests are correlated. The null hypothesis of no correlation can be rejected. A correlation of this degree might be a rough indication of group trends if used with other measures, but it would seem to have no predictive value for individual measures. It can probably be said, in general, that if an individual does very poorly on the sentence lists he would tend to do poorly on the word lists. It would seem that a correlation at this level would indicate that these two tests measure only somewhat similar--certainly not the same--kind of behavior.

Since the confidence interval for the *r* of C.I.D. Auditory Test W-22 has as its lower limit an *r* of .92, the null hypothesis cannot be rejected even though the reliability coefficient for this particular sample is .95.

The reliability coefficient computed for C.I.D. Everyday Speech is .98 with a .05 confidence interval from .97 to .99, which allows rejection of the null hypothesis stating the reliability will be less than .95.

Although both reliability coefficients are extremely

high, the sentence lists are significantly more reliable than the word lists as indicated by the .05 confidence intervals of their respective reliability coefficients. Because of this the null hypothesis stating that the reliability for C.I.D. Everyday Speech lists is less than or equal to the reliability of C.I.D. Auditory Test W-22 is rejected.

CHAPTER V

SUMMARY AND CONCLUSIONS

Summary

Determination of how an individual's hearing is functioning in everyday conversational situations plays an important part in current audiometric hearing evaluations. At the present time, word lists are used to determine the ability to hear and understand speech. Since we do not usually speak in single words, it was thought that a measure more closely approximating everyday conversation would be a more reliable means of testing the function of hearing in social situations. One way to approximate normal conversational speech situations more closely would be to use a test which utilizes sentences as the test material.

The C.I.D. Everyday Speech lists are a group of 100 sentences divided into lists of ten sentences each. These sentences were constructed so as to be representative of everyday speech. An articulation curve has been constructed for a recording of the sentence lists.

The purpose of this study was to compare these sentence lists with the PB word lists currently being used as a test of auditory discrimination. A low positive correlation was found between the two tests. High reliability

coefficients were found for the PB word lists and for the sentence lists. These were .95 and .98 respectively.

Conclusions

The low positive correlation between the two tests as a whole seems to indicate that, for the most part, these tests are not measuring the same variable. This would seem to supply a reason for the utilization of the sentence lists in audiometric hearing evaluation. If the two tests were measuring the same variable in the same way, the only reason for advocating the use of the sentence lists over the use of the PB word lists would be that the sentences can be administered more quickly than the PB word lists.

The sentence lists seem to have high "face validity." The reliability coefficient obtained suggests that there is probably little variation in results from form to form of this test. C.I.D. Everyday Speech was not constructed as a diagnostic measure, but rather is supposed to be a test of social functioning of hearing. It would appear to be a more valid method of testing social functioning in everyday situations than a test involving a single word presentation as the test material.

The previous discussion does not advocate the elimination of PB word lists in audiometric hearing evaluation. The purpose of the two tests seems to be different. According to Hirsh, C.I.D. Auditory Test W-22 is derived from a test with a primary purpose of diagnosis of hearing disorders.

The sentence lists are not in basic purpose a diagnostic test.

Implications for Further Research

This study utilized all normally hearing subjects. Noise was presented during the test administration to reduce scores. It would be interesting to learn how individuals with pathological hearing would score on the sentence lists. Noise is only one type of distortion. Hearing loss causes many other types.

During the course of this investigation many questions have arisen which warrant further research. Some of the questions are as follows:

1. Can the Social Adequacy Index be applied to scores obtained from use of these sentences?
2. Can these sentences be employed as a valid and reliable measure in hearing evaluations and in hearing aid evaluations?
3. Would using other distortion factors or pathological ears affect the results or reliability of these lists in any way?

BIBLIOGRAPHY

Books

Blalock, Hubert M. Social Statistics. New York: McGraw-Hill Book Company, Inc., 1960.

Davis, Hallowell, and Silverman, S. Richard. Hearing and Deafness. New York: Holt, Rinehart and Winston, Inc., 1960.

Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Company, Inc., 1959.

Greene, Edward B. Measurements of Human Behavior. New York: The Odyssey Press, 1952.

Hirsh, Ira J. The Measurement of Hearing. New York: McGraw-Hill Book Company, Inc., 1952.

Newby, Hayes A. Audiology. New York: Appleton-Century-Crofts, Inc., 1964.

Watson, Leland A., and Tolan, Thomas. Hearing Tests and Hearing Instruments. Baltimore: The Williams and Wilkins Company, 1949.

Periodicals

Campanelli, P. A. "Measurement of IntraList Stability of Four PAL Word Lists," Journal of Auditory Research, II (1962), pp. 50-55.

Davis, Hallowell. "The Articulation Area and the SAI for Hearing," Laryngoscope, LVIII (1948), pp. 761-78.

Egan, James P. "Articulation Testing Methods," Laryngoscope, LVIII (1948), pp. 955-91.

Elpern, B. S. "The Relative Stability of Half-List and Full-List Discrimination Tests," Laryngoscope, LXXI (1961), pp. 30-36.

Falconer, G. A., and Davis, H. "The Intelligibility of Connected Discourse as a Test for the 'Threshold of Speech,'" Laryngoscope, LVII (1947), pp. 581-95.

Fletcher, H., and Steinberg, J. C. "Articulation Testing Methods," Bell System Technical Journal, VIII (1929), pp. 806-54.

Grubb, Patti. "A Phonemic Analysis of Half-List Speech Discrimination Tests," Journal of Speech and Hearing Research, VI (1963), pp. 271-75.

Grubb, Patti. "Some Considerations in the Use of Half-List Speech Discrimination Tests," Journal of Speech and Hearing Research, VI (1963), pp. 294-97.

Hirsh, Ira J., Davis, Hallowell, Silverman, S. Richard, Reynolds, Elizabeth G., Eldert, Elizabeth, and Benson, Robert W. "Development of Materials for Speech Audiometry," Journal of Speech and Hearing Disorders, XVII (1952), pp. 321-37.

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

Miller, G. A., Heise, G., and Lichten, W. "The Intelligibility of Speech as a Function of the Context of the Test Materials," Journal of Experimental Psychology, XLI (1951), pp. 329-35.

Shutts, Edwin R., Burke, Kenneth S., and Creston, James E. "Derivation of Twenty-Five-Word PB Lists," Journal of Speech and Hearing Disorders, XXIX (1964), pp. 442-47.

Silverman, S. Richard, and Hirsh, Ira J. "Problems Related to the Use of Speech in Clinical Audiometry," Annals of Otology, Rhinology and Laryngology, LXIV (1964), pp. 1234-44.

Tobias, Jerry V. "On Phonemic Analysis of Speech Discrimination Tests," Journal of Speech and Hearing Research, VII (1964), pp. 98-100.

Zerlin, Stanley, and Urban, Beth. "The Evaluation of Contextual Speech Materials," Journal of Speech and Hearing Research, VI (1963), pp. 291-93.

Unpublished Materials

Johnson, Kenneth R. "Construction of an Articulation Curve for Recorded Sentences." Unpublished Master's thesis, Department of Speech, Michigan State University, 1965.

APPENDICES

APPENDIX A

Criteria for Sampling Everyday Speech as Suggested by CHABA[1]

1. The level should be specified in terms of relative frequency, age level, or educational level. There are numerous sources. The level should be of high frequency. The words should be common so that the test in no part depends upon vocabulary. The words should not be selected informally on the basis of personal estimate, but should appear in some specific fist [sic].

2. Within these objective limits, the vocabulary range should be fully exploited so that as many different words occur as possible.

3. Proper names and proper nouns should be excluded. They are unnecessary and unpredictable as to effect on validity.

4. Word length, measured in syllables, should be controlled. The ultimate test as a whole should have a distribution of good fit to the distribution of the vocabulary pool specified.

5. In the matter of syllable stress, in so far as this is inherent in words, free variation is suggested. Patterns and unusual departures from live speech should be avoided.

6. Contractions should be used freely and frequently. As a principle, they should be used whenever possible.

7. The ultimate test as a whole should have a phonetic frequency distribution that does not differ significantly from that of language, and this control should be demonstrated objectively by comparison to an existent criterion.

---

[1]S. Richard Silverman and Ira J. Hirsh, "Problems Related to the Use of Speech Audiometry," Annals of Otology, Rhinology, and Laryngology, LXIV (1955), pp. 1242-43.

Sentence Structure

1. The phonetic structure of a given sentence should be as such to avoid "loading," or unnaturally high frequency of occurrence of any one element, such as characterizes certain tongue-twisters. Within each sentence, the phonetic distribution should be at random.

2. To avoid testing memory span, an upper limit of sentence length should be set at 12 words. The lower limit should be fixed at two words, considering that many sentences are of this length and that one-word sentences should be avoided as duplicative of word tests. The distribution over this range, expressed in proportional parts, should be as follows:

Sentence length	parts
2 - 4	1
5 - 9	2
10 - 12	1

3. On the grounds that it will increase variety and interest, has face validity, and may be important (although the latter we do not know), the sentence form should be controlled as follows:

Sentence form	parts
Declarative	8
Imperative	1
Rising Interrogative	1/2
Falling Interrogative	1/2

This distribution appears not to depart far from that of American English in general.

4. Grammatical structure should vary freely and widely, and should avoid stereotyped forms.

5. Common, non-slang idioms should be used freely and it is desirable that they be numerous.

6. Redundancy should be high. An important aspect of validity is inference of unheard or incompletely heard material from fragments. In other words, it would not be good practice to build items all of which demand that every word be heard.

7. Sentence content should be appropriate for adults.

8. Levels of abstraction should be low to avoid the factors of intelligence, etc.

APPENDIX B

C.I.D. EVERYDAY SPEECH

Test Sentences[1]

List A

1. Walking's my favorite exercise.
2. Here's a nice quiet place to rest.
3. Our janitor sweeps the floors every night.
4. It would be much easier if everyone would help.
5. Good morning.
6. Open your window before you go to bed!
7. Do you think that she should stay out so late?
8. How do you feel about changing the time when we begin work?
9. Here we go.
10. Move out of the way!

List B

1. The water's too cold for swimming.
2. Why should I get up so early in the morning?
3. Here are your shoes.
4. It's raining.
5. Where are you going?
6. Come here when I call you!
7. Don't try to get out of it this time!
8. Should we let little children go to the movies by themselves?
9. There isn't enough paint to finish the room.
10. Do you want an egg for breakfast?

List C

1. Everybody should brush his teeth after meals.
2. Everything's all right.
3. Don't use up all the paper when you write your letter.

---

[1]Hallowell Davis and S. Richard Silverman, Hearing and Deafness (New York: Holt, Rinehart and Winston, Inc., 1961), pp. 549-52.

4. That's right.
5. People ought to see a doctor once a year.
6. Those windows are so dirty I can't see anything outside.
7. Pass the bread and butter, please!
8. Don't forget to pay your bill before the first of the month.
9. Don't let the dog out of the house!
10. There's a good ballgame this afternoon.

List D

1. It's time to go.
2. If you don't want these old magazines, throw them out.
3. Do you want to wash up?
4. It's a real dark night so watch your driving.
5. I'll carry the package for you.
6. Did you forget to shut off the water?
7. Fishing in a mountain stream is my idea of a good time.
8. Fathers spend more time with their children than they used to.
9. Be careful not to break your glasses!
10. I'm sorry.

List E

1. You can catch the bus across the street.
2. Call her on the phone and tell her the news.
3. I'll catch up with you later.
4. I'll think it over.
5. I don't want to go to the movies tonight.
6. If your tooth hurts that much you ought to see a dentist.
7. Put that cookie back in the box!
8. Stop fooling around!
9. Time's up.
10. How do you spell your name?

List F

1. Music always cheers me up.
2. My brother's in town for a short while on business.
3. We live a few miles from the main road.
4. This suit needs to go to the cleaners.
5. They ate enough green apples to make them sick for a week.
6. Where have you been all this time?
7. Have you been working hard lately?

8. There's not enough room in the kitchen for a new table.
9. Where is he?
10. Look out!

List G

1. I'll see you right after lunch.
2. See you later.
3. White shoes are awful to keep clean.
4. Stand there and don't move until I tell you!
5. There's a big piece of cake left over from dinner.
6. Wait for me at the corner in front of the drugstore.
7. It's no trouble at all.
8. Hurry up!
9. The morning paper didn't say anything about rain this afternoon or tonight.
10. The phone call's for you.

List H

1. Believe me!
2. Let's get a cup of coffee.
3. Let's get out of here before it's too late.
4. I hate driving at night.
5. There was water in the cellar after that heavy rain yesterday.
6. She'll only be gone a few minutes.
7. How do you know?
8. Children like candy.
9. If we don't get rain soon, we'll have no grass.
10. They're not listed in the new phone book.

List I

1. Where can I find a place to park?
2. I like those big red apples we always get in the fall.
3. You'll get fat eating candy.
4. The show's over.
5. Why don't they paint their walls some other color?
6. What's new?
7. What are you hiding under your coat?
8. How come I should always be the one to go first?
9. I'll take sugar and cream in my coffee.
10. Wait just a minute!

List J

1. Breakfast is ready.

2. I don't know what's wrong with the car, but it won't start.
3. It sure takes a sharp knife to cut this meat.
4. I haven't read a newspaper since we bought a television set.
5. Weeds are spoiling the yard.
6. Call me a little later!
7. Do you have change for a five-dollar bill?
8. How are you?
9. I'd like some ice cream with my pie.
10. I don't think I'll have any dessert.

APPENDIX C

C.I.D. AUDITORY TEST W-22[1]

List 1A

an	you (ewe)
yard	as
carve	wet
us	chew
day	see (sea)
toe	deaf
felt	them
stove	give
hunt	true
ran	isle (aisle)
knees	or (oar)
not (knot)	law
mew	me
low	none (nun)
owl	jam
it	poor
she	him
high	skin
there (their)	east
earn (urn)	thing
twins	dad
could	up
what	bells
bathe	wire
ace	ache

[1] Auditory Test W-22 (St. Louis, Missouri: Central Institute for the Deaf).

List 2A

yore (your)
bin (been)
way
chest
then
ease
smart
gave
pew
ice
odd
knee
move
now
jaw
one (won)
hit
send
else
tare (tear)
does
too (two, to)
cap
with
air (heir)
and
young
cars
tree
dumb
that
die (dye)
show
hurt
own
key
oak
new (knew)
live (verb)
off
ill
rooms
ham
star
ear
thin
flat
well
by (buy)
ail (ale)

List 3A

bill	aim
add (ad)	when
west	book
cute	tie
start	do
ears	hand
tan	end
nest	shove
say	have
is	owes
out	jar
lie (lye)	no (know)
three	may
oil	knit
king	on
pie	if
he	raw
smooth	glove
farm	ten
this	dull
done (dun)	though
use (yews)	chair
camp	we
wool	ate (eight)
are	year

List 4A

all (awl)	darn
wood (would)	art
at	will
where	dust
chin	toy
they	aid
dolls	than
so (sew)	eyes (ayes)
nuts	shoe
ought (aught)	his
in (inn)	our (hour)
net	men
my	near
leave	few
of	jump
hang	pale (pail)
save	go
ear	stiff
tea (tee)	can
cook	through
tin	clothes
bread (bred)	who
why	bee (be)
arm	yes
yet	am

APPENDIX D

TEST PRESENTATION ORDER

Subject Number	C.I.D. Auditory Test W-22 Form	C.I.D. Everyday Speech Sentence List
1	1A	A
2	2A	B
3	3A	C
4	4A	D
5	1A	E
6	2A	F
7	3A	G
8	4A	H
9	1A	I
10	2A	J
11	3A	A
12	4A	B
13	1A	C
14	2A	D
15	3A	E
16	4A	F
17	1A	G
18	2A	H
19	3A	I
20	4A	J
21	1A	A
22	2A	B
23	3A	C
24	4A	D
25	1A	E
26	2A	F
27	3A	G
28	4A	H
29	1A	I
30	2A	J
31	3A	A
32	4A	B
33	1A	C
34	2A	D
35	3A	E
36	4A	F
37	1A	G
38	2A	H
39	3A	I
40	4A	J

APPENDIX E

TEST FORM

C.I.D. Auditory Test W-22

Subject Number_____ Date_____

Name_____ Test Sequence_____ _____

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

TEST FORM

C.I.D. Everyday Speech

Subject Number_____ Date_____

Name_____ Test Sequence_____ _____

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

APPENDIX F

TEST SCORES

	Full Test Scores		Half Test Scores			
			W-22		Sentences	
	W-22	Sentences	Odd	Even	Odd	Even
1	34	46	20	48	48	40
2	38	34	32	44	36	32
3	32	52	36	28	52	52
4	50	68	52	48	68	68
5	30	26	32	28	24	28
6	32	30	20	44	24	36
7	42	60	44	40	60	60
8	44	60	44	44	56	64
9	42	40	44	40	40	40
10	28	34	28	28	32	36
11	38	74	40	36	76	72
12	40	64	48	32	60	68
13	44	60	36	52	64	56
14	32	22	24	40	24	20
15	50	54	52	48	52	56
16	52	56	48	56	56	56
17	34	58	36	32	56	60
18	36	22	32	40	20	24
19	60	54	64	56	68	40
20	22	24	28	16	20	28
21	52	64	60	44	64	64
22	36	44	36	36	44	44
23	40	58	40	40	64	52
24	48	48	52	44	52	44
25	48	56	48	48	56	56
26	48	68	48	48	64	72
27	62	60	60	64	60	60
28	40	68	44	36	68	68
29	36	52	24	48	52	52
30	34	46	32	36	44	48
31	42	76	44	40	72	80
32	40	64	32	48	64	64
33	36	56	36	36	56	56
34	44	30	52	36	28	32
35	36	40	36	36	40	40
36	50	38	36	64	36	40
37	32	58	32	32	56	60
38	30	80	24	36	84	76
39	30	44	20	40	52	36
40	24	20	20	28	20	20