1972

Comparison of the Goldman-Fristoe-Woodcock and Modified Rhyme Tests of auditory discrimination

James Edward Mangan
The University of Montana

COMPARISON OF THE GOLDMAN-FRISTOE-WOODCOCK AND MODIFIED RHYME TESTS OF AUDITORY DISCRIMINATION

By

James E. Mangan

B.A., Carroll College, 1966

Presented in partial fulfillment of the requirements
for the degree of

Master of Arts
UNIVERSITY OF MONTANA

1972

Approved by:

B. B. Chaney
Chairman, Board of Examiners

John M. Stewart
Dean, Graduate School

Aug 21, 1972
Date

8-21-72

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Auditory discrimination, or more specifically, the ability to distinguish between closely related speech sounds, has long held the interest of speech pathologists and audiologists. This phenomenon has been studied from many vantage points. Wood (1971), in a review of the literature, considered many different aspects of audition, including the nature of the auditory signal, the psychologic set of the listener and the methods of assessing informational transfer. In addition, studies have been conducted using speech reception threshold tests and phonemic discrimination tests. Both verbal and motor responses signifying comprehension have been reported. In general, studies of auditory discrimination in adults have shown that the intensity, frequency and directionality of the signal; the timing, content and number of extraneous background signals; and the commitment of the listener to attend to a shift from one simultaneous signal source to another under differing cue conditions are just a few of the major variables that combine to make auditory discrimination a complex psychological process.

Some studies have focused on etiological factors which are known to interfere with the ability to discriminate between and among auditory signals. Goldstein (1948), for example, hypothesized that the central nervous system must be intact in order for the subject to successfully organize complex sets of incoming auditory information. This conclusion was supported by the research of Winchester and Hartman (1955), who studied the figure-ground responses of brain-injured adults. On the basis of that study, the investigators reported that subjects with central nervous system involvement had performed significantly poorer than normal subjects on auditory figure-ground discrimination tasks in which speech was presented against a white noise background. Strauss and Lehtinen (1947), demonstrated that brain-damaged retardates were distracted more than non-brain damaged retardates were by auditory background stimulation. Strothers (1963), discussed this point in fair detail in terms of evaluation and habilitative management of children with neurological impairment.

A vital consideration in the exploration of any phenomenon is the nature of the measuring instrument used. The phenomenon is never known in and of itself, but only through the "lens," both clarifying and distorting, of the measuring device. A variety of tests have been devised for use in the exploration of auditory discrimination and these tests vary in form and content.

The use of speech testing for possible hearing impairment undoubtedly preceded the use of discrete tones by many years, but the quantification of speech audiometric tests has been fairly recent. The first systematic use of speech testing was not concerned with the hearing impaired population, but rather was used to evaluate the speech-reproducing capabilities of various sound transmitting and amplifying systems. Berger (1971), cited studies that were conducted as early as the 1920's and that proposed a method of calculating the efficiency of telephone sound transmitting equipment using nonsense syllables consisting of various consonants followed by the vowel /i/. A group of speakers called these nonsense syllables over communication systems, and in this way, the consonant sounds could be evaluated as to how well or poorly they were reproduced by the systems under study. A refinement of the syllable choice was made wherein initial, medial and final sounds in syllables were obtained by a random selection, and the combination of consonant-vowel was varied in a systematic manner. The result was 174 different lists with 50 nonsense syllables in each. These lists became known as the Standard Articulation Test. Another study employed two different speakers and seven or eight listeners at a time to evaluate telephone systems, using the Standard Articulation Test. The efficiency of each test instrument was calculated by assigning to it the percentage of syllables correctly heard. Thus, arose the "Percent Score of

Articulation." An advantage of nonsense syllables in testing speech discrimination is their lack of meaning, and thus, the listener's vocabulary is not a variable. But this lack of meaning in turn can also be a disadvantage since the listener does not normally need to identify meaningless stimuli. Persons trained in such listening tasks manipulate nonsense syllables well, but those untrained generally do very poorly. For those reasons, it was recommended that mono-syllabic words be used for the testing of the deaf as early as 1929.

The "Grandfather" of auditory discrimination tests is that constructed by Travis and Rasmus (1931) for their pioneer investigation of the subject. It is, without a doubt, the most extensive ever used in these studies. Every sound in the English language was compared with every other sound and with itself. The sounds were compared on the basis of alikeness or a difference in quality only. The test consists of 366 pairs of speech sounds, 300 pairs being consonants and 66 pairs vowel combinations. In its original form, the test has been used only in this first experiment, and in the much more carefully controlled one by Hall (1938).

The task of making 366 comparisons was quickly recognized as overwhelming even for an adult. Nonetheless, the Travis-Rasmus Test has served as the model for most tests, at least in general form, if not in specific content. The first standardization of this type of test was that done by Templin (1943), in presenting a new test that was based on the Travis-Rasmus model, but which contained only 70 pairs of nonsense syllables.

Another method of approaching the problem of measurement is that of asking the subject to distinguish an incorrect sound or word in the context of a sentence. Hall's test (1938) involves the use of coined words in a meaningful phrase. Mase's test (1946) presents a word which is similar in sound to the correct word, but quite different in meaning, (e.g. "shoe" substituted for "chew"). In this latter study, it was found that the subjects as a group performed less well on a sentence test than a word pair discrimination test which used the same words.

Another group of tests may be termed "individual," since each subject is given a test whose content is based on his own articulation errors. Farquhar's test (1961), required the subject to clap hands each time he heard the examiner utter the correct form of his misarticulated sound. The test was intended for use with kindergarten children. Spriestersbach and Curtis, (1951), investigated discrimination of a speech sound such as /s/ in various contexts by imitating the child's misarticulation of that sound as closely as possible in one of three utterances of a word. The child's task was to indicate the correct production.

There is another test to be noted, which, again, is a paired comparison test, but one which seeks to come to terms with the problem of the abstract nature of the Travis-Rasmus model. As Templin noted (1957), the original test "demands considerable intellectual development." This derives both from the necessity to understand the concepts of "same and different" and the difficulty presented by the use of nonsense syllables. Reducing the number of pairs served to bring the examination to a manageable size, but it did not eliminate the problem, as is evident from the positive relationship between the discrimination scores and intelligence test scores. In the Wepman Auditory Discrimination Test (1958), the method of presenting sound pairs which differ in only one sound, or are identical, is retained, but word pairs (40 in number) rather than nonsense syllables are used as the stimuli. The abstract character of the test is thereby reduced, as is the difficulty of the test, since a difference in meaning rather than merely in sound value is created by the use of the real words. Correlation with intelligence test results was found to be 0.32 (Spearman Rank Order Correlation).

A test which has been used clinically quite often is the Phonetically Balanced Word List (PB-50). These lists, 50 words each, of reasonably familiar monosyllables, were developed at the Harvard Psycho-Acoustic Laboratory for wartime research on equipment for communication. In this test, nearly all the phonemes of the English language are represented with appropriate frequency in every list of 50 words. If the initial consonants and vowels are considered, the frequencies of occurrence of the various sounds is fairly representative of English speech as a whole.

In a recorded version of this test, the speaker spoke the test item number at a set intensity by peaking this as a carrier phrase on a vu-meter; the key, or test word, following the item number was spoken at an intensity in relation to the carrier phrase at which it might normally occur. The actual test monosyllables did not all occur at the same level, although it has long been known that small variations in intensity level may cause rather large variation in discrimination for speech.

In response to criticisms of word difficulty and several other factors, a group at the Central Institute for the Deaf amended the PB-50 lists and produced the familiar W-22 recordings. The speaker followed the procedure of saying a carrier phrase ("you will say ___") at a set level and allowing the key word to fall at some presumably natural level relative to the carrier phrase. The W-22 lists consist of 120 words from the PB-50's plus 80 new words. The 200 test words are divided into four basic lists of 50 words each, and these in turn are scrambled in six orders each. One further modification in the use of the W-22 recordings and lists has been the division of the standard lists into half lists. The user of the recordings or lists simply employs 25 rather than all 50 word items; using 50 words allows a handy two percent weight for each item correctly identified; using only 25 words means each word now has a value of four percent, with the possibility that any single error by the listener or examiner now receives twice the error score as he did previously and as a result the variability will be higher.

The Picture Auditory Discrimination test forms a quite different instrument from those already described, in that the subject is asked to respond to stimuli in two sense modalities--the auditory and the visual. Pictures are presented to the subject and he must tell whether or not the words represented by the pictures sound alike, or which picture is named by the tester. Two picture tests that have been used in experimental investigations (Schiefelbusch and Lindsey, 1938; Templin, 1957), vary both in number of pictures presented in time and in the method of administration.

A picture test that has recently been developed, and that is presently receiving a considerable amount of clinical attention is the Goldman-Fristoe-Woodcock Test of Auditory Discrimination. This test is designed (according to its publishers) to provide a measure of speech-sound discriminability, relatively unconfounded by other factors. It provides a measure of auditory discrimination under ideal listening conditions plus a comparative measure of auditory discrimination in the presence of controlled background noises (cafeteria noises). The test consists of three parts which are presented binaurally. The first is the training procedure, which familiarized the subjects with the word-picture associations to be used during the two subtests, and permits the examiner to establish the presence of these associations. The second part is the quiet subtest, and the third part is the noise subtest. Separate norms are provided for each subtest from ages 3 years 8 months to 70 years and over.

Fairbanks, in 1958, described a method for testing phonemic differentiation of speech reception. He developed five 50 item tests which he considered to be reasonably comparable. The test items were drawn from 250 common monosyllabic words. There were 50 sets of five rhyming words. One item from each of the 50 sets appeared on a given test form. Words for the tests are provided in randomized order.

On the response sheet for the Fairbanks' Rhyme Test, there are five columns of ten"stems." A stem is the rhyming portion of the five-word sets. The rhyming words differ only in the initial consonant. So when the subject hears a word, he is asked to put the proper letter in the space preceding the stem. Response sheets would consist of items such as ot, -ay, -op. Copies of the same response sheet would be used for all five test forms.

An additional list, Rhyme Test F, is loaded with familiar words, and might be suitable for testing a child or a semi-literate adult.

In 1963, and again in 1965, House, et.al., published articles concerning the development of the Modified Rhyme Test (MRT). The test itself has a multiple choice format; the subject is asked to draw a line through the correct response. Each item contains five foil words in addition to the correct responses. On a test, there are fifty rectangles which contain the six words of each item. An illustration of part of a test appears below:

mop	1	shop	lane	2	lame	beach	3	beat	sang	4	hang
top		hop	lace		lay	beam		beak	gang		bang
cop		pop	lake		late	bead		beam	rang		fang

As a rule, the words were consonant-vowel-consonant (CVC), though some were CV or VC. Each of six rhyming words appeared on a test form. Only a single initial or final consonant is varied; the remainder of the word consistent with its foils. No attempt was made to phonetically balance the lists, though there was an attempt to have the same speech sound representation in each form. The tests were presented at six different signal to noise ratios.

In 1968, Kruel, et.al., attempted to adapt the MRT to the needs of the clinical audiologist. The MRT materials of House, et.al., were only slightly modified. Kruel, et.al., recorded the MRT tests onto six tapes which are now available for use as a clinical tool. The speakers on these tapes are two males and one female. Each speaker presents a different form of the test. There are two tapes for each of the three speakers and there are four test lists on each tape. The four lists are presented at different S/N ratios using white noise to produce different reference levels of speech discrimination. The order of test conditions for each tape is as follows: +30S/N level, P83, P75 and +30S/N. Each tape includes a calibration tone and noise, speaker and list identification, suitable instructions for the listeners and practice items. A person with normal hearing is expected to receive a score of 96% at the +30S/N level, a score of 83% at P83 and a score of 75% at P75.

The Goldman-Fristoe-Woodcock Test of Auditory Discrimination is used in approximately 25% of the outpatient evaluations at the University of Montana. It is reportedly used quite frequently at other facilities throughout the state and elsewhere. In view of the frequent and growing clinical use of the Goldman-Fristoe-Woodcock, the experimenter felt that there was a need for a study in which the following question could be answered: Will the Goldman-Fristoe-Woodcock Test of Auditory Discrimination, a short, easily administered test, give a clinician as much information concerning speech discrimination as the longer, more sophisticated Modified Rhyme Test?

CHAPTER II

PROCEDURE

Statement of the Problem

This investigation was intended to determine if there is a correlation between the individual scores of the Goldman-Fristoe-Woodcock and the Modified Rhyme tests of auditory discrimination.

Upon studying the normative data of both tests, it became apparent that a strong similarity existed in the means of both tests in both quiet and noise environments (Appendix A). The hypothesis to be tested was: There is no correlation between the two tests when individual scores are compared.

The experimental variables in this study are:
(a) Independent: exposure to a testing situation in which the Goldman-Fristoe-Woodcock Test of Auditory Discrimination and sections of the Modified Rhyme Test are presented;
(b) Dependent: a percentage score achieved by the listeners on these tests.

Materials and Procedures Used in the Experiment

Two test tapes were constructed using a Viking 235 Tape Duplicator. Both the quiet and noise (+9S/N ratio) subtests of the Goldman-Fristoe-Woodcock Test of Auditory Discrimination (GFW) were recorded as were the following portions of the Modified Rhyme Test (MRT): the "D" list which has a +30S/N ratio and the "B" list which has a +10S/N ratio. The "D" list thus approximates the quiet subtest of the GFW and the "B" list approximates the noise subtest. Test Tape 1 will first present the GFW, then sections of the MRT. Test Tape 2 will present the MRT first, followed by the GFW.

All subjects were tested in the Speech and Hearing Clinic at the University of Montana. A total of 40 subjects were selected from both male and female volunteers ranging in age from 19 to 61 years, from the Missoula, Montana area. The subjects varied in their educational experience (8th grade to college graduate) and in their hearing acuity. Out of the sample tested, 22 individuals had a hearing loss greater than 20 dB at one or more of the test frequencies (Table I).

Each subject was given a pure tone test in the right ear at the frequencies of 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz. At the time of selection, each subject received a number from 1 to 40. These numbers were then used to separate the subjects into two groups, the odd numbers comprising one group and the even numbers comprising the second group.

Each subject participated individually in the experiment. Upon entering the sound-treated room, the subject received a set of instructions concerning the experiment (Appendix B). All the odd numbered subjects (1, 3, 5, etc.) listened to the GFW test first, then the MRT. Those with

TABLE I

A Distribution of the Hearing Losses As a Function of Intensity and Frequency

Frequencies in Hertz

Hearing Level in Decibels	2000	3000	4000	6000	8000
Less Than 20	36	35	26	21	36
20	1		4	6	1
25					
30	1		2	2	
35					
40				1	
45		1	1		
50	1	1	1	3	1
55			1		
60		1	1	1	
65	1			2	1
70			1		
75					
80		2	2	2	
85					
90			1	1	
95				1	
100					
Total	40	40	40	40	40

even numbers (2, 4, 6, etc.) listened to the MRT first followed by the GFW. Both tests were presented at 40 dB above the pure tone averages for the best two of three frequencies 500, 1000 and 2000 Hz. It should be noted that the GFW was originally designed to be presented binaurally, whereas the MRT is a monaural test. To eliminate the variable of binaural vs monaural presentation, both tests were arbitrarily administered to the right ear. A 1000 Hz tone on the MRT test tape was used to calibrate the audiometer prior to the presentation of the test tapes.

The test tapes were presented to the subjects using the Sony tape recorder (Model TC-366) and a Grason-Stadler Audiometer (Model 1701) with Grason-Stadler ear phones (Model TDH 49-10Z) in a sound treated IAC room at 7½ i.p.s. The experimenter was seated next to and facing the subject in order to present the plates of the GFW. The experimenter also wore ear phones in order to monitor the testing procedures. When the MRT was presented, the examiner was not present in the test room.

Listeners responses: When the GFW was presented, an Easel Kit was used which contained plates with four pictures each. One of the pictures would represent the word spoken on the test tape and the subject was required to point to the appropriate picture. When the MRT was presented, the subject was presented with a score sheet which has a multiple choice format. The subject was asked to draw a line through the correct response. Each item contains five foil words in addition to the correct response. There are 50 rectangles which contain the six words of each item (Appendix C). The responses of the subjects were marked on the appropriate score sheets and the raw scores were converted into percentage scores for statistical analysis.

Experimental Design

The basis of this study was to determine the degree of correlation that exists between the GFW and the MRT tests of auditory discrimination. The subjects were randomly divided into two groups in order to counterbalance the learning effect. The Pearson Product Moment Test of Correlation was used and a .05 significance level was chosen. The obtained correlation was tested by means of a t test to determine if it was significantly different than zero at this level.

CHAPTER III

RESULTS AND DISCUSSION

The results obtained in this investigation were evaluated by means of a Pearson Product Moment Test of Correlation. The major concern of this investigation was to determine the degree of correlation between the individual scores of the Goldman-Fristoe-Woodcock and the Modified Rhyme Tests of Auditory Discrimination. A summary of an analysis of the data is presented in Table II. The order effect was not significant and a summary of the statistics for that effect is presented in Table III.

In view of the fact that more than half of the subjects tested had a hearing loss greater than 20 dB at one or more of the test frequencies, the subjects were again separated into two groups. The first group (Group I) did not have a hearing loss greater than 20 dB, and the second group (Group II) did have such a loss. The data from these two groups was then analyzed and a summary of the statistics is presented in Table IV.

TABLE II

Means, Variances, Standard Deviations
and Correlation Coefficients
for the GFW and MRT Tests of Auditory Discrimination

	GFW	MRT	r	r^2	t
Quiet Subtest			.7206*	.5192	6.4035**
$\bar{X}$	94	88			
σ^2	39.6	38.8			
σ	6.25	6.23			
Noise Subtest			.4611*	.2126	3.2032**
$\bar{X}$	76.6	77			
σ^2	88.6	38.8			
σ	9.41	6.63			
Composite Score			.7000*	.4900	6.042**
$\bar{X}$	85.3	82.6			
σ^2	51.7	31.4			
σ	7.19	5.61			

* The correlations obtained in this investigation were significant beyond the .05 significance level.

** The obtained correlations were found to be significantly different from zero at the .05 level.

TABLE III

A Summary of Statistics for the Order Effect

	Quiet Subtests				Noise Subtests			
	$\bar{X}$	σ^2	σ	t	$\bar{X}$	σ^2	σ	t
GFW								
Order I	94.9	44.5	6.67	.997*	74.4	108.5	10.41	.742
Order II	92.9	31.8	5.64		77.7	66.0	8.1	
MRT								
Order I	87.8	37.9	6.16	.346*	76.2	49.0	7.0	.751
Order II	88.5	39.5	6.2		77.8	37.5	6.1	

* Although the t scores for the quiet subtests appear to differ markedly, given the similar means and variances, both t values are still far from a significant value of 2.093 at the 5% level of significance.

TABLE IV

Means, Variances, Standard Deviations
And Correlation Coefficients
For Subjects as a Function of Their Hearing Losses

	Group I (n=18) Hearing Loss Less Than 20 dB			Group II (n=22) Hearing Loss Greater Than 20 dB		
	GFW	MRT	r	GFW	MRT	r
Quiet Subtest			.327			.784
$\bar{X}$	95.7	90.0		92.4	86.6	
σ^2	11.3	22.2		56.7	47.4	
σ	3.3	4.71		7.5	6.8	
Noise Subtest			.315			.606
$\bar{X}$	80.3	77.0		77.3	77.0	
σ^2	25.2	50.3		120.7	38.8	
σ	5.02	7.09		10.9	6.2	

Discussion

The method employed to estimate the equivalance of tests is determined to some extent by the use planned for the test. When a test is used for evaluating speech intelligibility through a communication channel, equal means and variances from repeated presentations are probably satisfactory indicators of reliability. No assumption need be made that the test measures a variable on which the listeners differ. A test of speech discrimination, however, is presumed to measure some variable on which listeners really do differ. Such a test, therefore, should then discriminate among the listeners being tested according to their differing abilities.

The criteria necessary for tests to be considered as measuring the same thing in different listeners include not only equal means and variances, but a high correlation as well (Bell and Kreul, 1972), In many such cases, the least important consideration is similarity of means, with correlation being the more important comparison.

The basic concern of this investigation was to determine the degree of correlation between the individual scores of the Goldman-Fristoe-Woodcock and the Modified Rhyme tests of auditory discrimination. In reviewing the data, a strong similarity exists between both the means and variances of the comparable MRT and GFW subtests. A similar relationship exists for the composite scores of both tests. A relatively high correlation exists on the quiet subtests of both tests (.72), and on the composite scores of both tests (.70). On the noise subtests, a much lower correlation was attained (.46), which indicates more unshared than shared variance on those tests.

In a test-retest reliability study, the authors of the GFW reported that the GFW was administered to a sample of 17 preschool speech handicapped children at the Bill Wilkerson Hearing and Speech Center in Nashville, Tennessee. Two weeks later, this group was again administered the GFW. A test-retest correlation of .87 was obtained on the quiet subtest and a correlation of .81 was obtained on the noise subtest. One variable that this study did not account for, that may have influenced the correlation, was the learning effect.

Bell and Kreul (1972) reported, in a study concerning the reliability of the MRT, a correlation of .92 with the quiet subtest ("D" list) and a correlation of .57 with the noise subtest ("B" list). That study was conducted on two groups of subjects. The first group of subjects consisted of 14 young male and female college students with normal hearing.[1] The second group of subjects consisted of 12 patients who were obtained from the Palo Alto Veterans

---

[1]In that study, all listeners showed a 15 dB or better threshold hearing level (re: ISO-1964 audiometric zero) at 500, 1000, 2000, 3000 Hz, and 25 dB or better at 4000 and 6000 hZ.

Administration Audiology Clinic. These subjects seemed to have a noise-induced hearing loss, and exhibited a sharp drop in sensitivity above 3000 Hz, with a history of noise exposure. The low correlation shown on the retest data of that study may explain why such a low correlation was obtained in the present investigation on the noise subtests of the two measures.

The "D" list of the MRT was chosen for this experiment because it most closely approximated the signal to noise environment of the GFW, even though it had one of the lowest test-retest correlations (.57). Other subtests of the MRT however, revealed test-retest correlations that were considerably higher (up to .92). Thus, it would appear that the MRT may be a more sensitive instrument than perhaps the "D" list taken alone would suggest.

The correlation that was obtained from the subjects, who did not have a hearing loss at any of the test frequencies (Group I), was .327 for the quiet subtests and .315 for the noise subtests, again indicating more unshared than shared variance. Although the two means appear relatively close, the distribution of scores around the means is such that there is very little overlap. The low correlation, then, is not surprising. It thus appears that the two tests do not give similar estimates of auditory discrimination abilities in normal hearing listeners.

The correlations obtained from those subjects with hearing losses greater than 20 dB at one or more of the test frequencies (Group II), was .784 for the quiet subtest and .606 for the noise subtests. The relatively larger variances could indicate that the measuring instruments are sensitive enough to display varying auditory discrimination abilities among these listeners. The higher correlations would suggest that the two tests do so in a similar manner. If it is the listeners with hearing losses that are our primary clinical concern, then it becomes less critical whether these tests differentiate discrimination abilities among normal listeners. A rough interpretation of the foregoing is that the two tests give more similar results with hearing impaired populations than with normal listeners.

The test-retest reliability information and the results of this investigation indicate that one obtains similar estimates of discrimination ability when comparing the two tests as one does when each test is used separately.

It was apparent that the amount of variance that was shared between these two tests was between 21% and 52%. To some, this may appear to indicate that neither test is a precise enough instrument to be used clinically. Others may feel that this amount of shared variance is better than none, and thus could be considered acceptable when ease of administration and time factors are included in the choice of a test. It is recognized by this investigator that there currently appears to be a lack of precision in the measuring instruments that are available for testing auditory discriminations.

It should be noted that this investigation attempted to explore the degree of equivalance of the two measures and not the variable they were designed to measure. Numerous researchers have addressed themselves to this issue and it appears that this is an area worthy of further investigation. Once the variable under consideration (auditory discrimination) is more clearly identified and defined, perhaps a more precise and reliable measuring instrument will be devised. The fact that several new auditory discrimination tests have been produced within the past year, suggests that there is a widespread concern and need to find better measuring instruments.

If the test procedures, as outlined in this investigation, are followed, it appears that the GFW and MRT do, in fact, give similar results. Since the GFW is a short (7½ minute), easily administered test which can be used with individuals ranging in age from 3 years 8 months to 70 years and over, it may be preferred by many clinicians. The MRT requires that the individual be able to read, and it takes approximately 40 minutes to administer. In the opinion of this investigator, either test appears to be acceptable as a measure of auditory discrimination abilities as one can use considering the tests that are currently available. However, it is felt that there is a need for a test with higher internal reliability for a clinical population with varying hearing abilities. Hopefully, such a test can soon be made available to the clinician.

CHAPTER IV

SUMMARY AND CONCLUSIONS

An investigation was made to determine the degree of correlation between the individual scores of the Goldman-Fristoe-Woodcock and Modified Rhyme tests of auditory discrimination. The question of concern was: "Do the two tests give similar results when used on the same listeners?"

A total of 40 subjects participated in this experiment who varied in age, education and hearing acuity. Both the GFW and the MRT were presented monaurally to the right ear 40 dB above the pure tone averages for the best two of three frequencies 500, 1000, and 2000 Hz. Subjects were randomly divided into two groups in order to counterbalance the learning effect.

The results obtained were evaluated by means of a Pearson Product Moment Test of Correlation. A correlation was obtained for the quiet subtests, noise subtests and composite scores of both tests on hearing impaired and normally hearing listeners. A relatively high correlation was obtained on the quiet subtests of both tests. The noise subtests showed a lower correlation, indicating more unshared than shared variance on those subtests. The order effect was not significant and a strong similarity exists between the means and variance of both tests.

The test-retest reliability information from each test and the comparisons made in this investigation indicate that one obtains similar estimates of discrimination ability with either of the two tests, and the correlation is better with hearing impaired than with normally hearing listeners.

BIBLIOGRAPHY

Bell, D.W., Kreul, E.J., Modified Rhyme Test Reliability, (in press), Journal of Speech and Hearing Disorders, (1972).

Berger, K.W., Speech Audiometry (in) Audiological Assessment, (Ed. Rose, D.E.), Prentice-Hall, Inc., Englewood Cliffs, New Jersey, (1971).

Fairbanks, G., "Test of Phonemic Differentiation: The Rhyme Test," Journal of the Acoustical Society of America, 30, 596-600, (1958).

Farquhar, Mary S., "Prognostic Value of Imitative and Auditory Discrimination Tests," Journal of Speech and Hearing Disorders, 26, 342-347, (1961).

Goldstein, K., Language and Language Disturbances, New York: Grune and Stratton (1948).

Hall, Margaret C., "Auditory Factors in Functional Articulatory Speech Defects," Journal of Experimental Education, 7, 110-132, (1938).

House, A.S., Williams, C.E., Hecker, M.H.L., and Dryter, K.P., Psychoacoustic Speech Tests: A Modified Rhyme Test, U.S. Air Force Systems Command, Hanscon Field, Electronics System Division, Tech. Doc. Report SED-TDR-63-403, (June, 1963).

Kreul, E.J., Bell, D.W., and Nixon, J.C., "Factors Affecting Speech Discrimination Test Difficulty," Journal of Speech and Hearing Research, 12, 281-287, (1969).

Kreul, E.J., Nixon, J.C., Dryter, K.P., Bell, D.W., and Lang, J.S., "A Proposed Clinical Test of Speech Discrimination," Journal of Speech and Hearing Research, 11, 536-552, (1965).

Mase, P.J., Etiology of Articulatory Speech Defects, New York: Teachers College, Columbia Univ., Bureau of Publications, Contributions to Education, No. 921, (1946).

Schiefelbusch, R.L., and Lindsey, Mary J., "A New Test of Sound Discrimination," Journal of Speech and Hearing Disorders, 23, 153-159, (1958).

Spriestersbach, D.C., and Curtis, J.F., "Misarticulation and Discrimination of Speech Sounds," Quarterly Journal of Speech, 37, 483-491, (1951).

Templin, Mildred C., "A Study of Sound Discrimination Ability of Elementary School Pupils," Journal of Speech Disorders, 8, 127-132, (1943).

Templin, Mildred C., Certain Language Skills in Children, Minneapolis: Univ. Minn. Press (1957).

Travis, L.E., and Rasmus, Bessie, "The Speech Sound Discrimination Ability of Cases With Functional Disorders of Articulation," Quarterly Journal of Speech, 17, 217-226, (1931).

Wood, Nancy E., "Final Report," Auditory Perception in Children, Social Rehabilitation Services, Research Grant No. RD-25-75-5., (1971)

APPENDIX A

Both the GFW and the MRT were standardized using approximately 40 individuals from the same age group. Standardization means and standard deviations are listed below:

	$\bar{X}$	σ
GFW quiet	98.4	2.8
GFW noise	81.2	6.96
MRT quiet	98.1	2.42
MRT noise	88.2	3.88

APPENDIX B

Instructions and Introductory Statement Made by the Experimenter to the Subject.

My name is ______________. Please come in and be seated. You are about to take part in an experiment. When the experiment has been completed, you will receive a summary explaining the purpose and the results of the experiment.

You are about to listen to two tape recordings of a list of words. The first tape that you hear will require that you point to a picture which represents the word that you will hear. I will say a word. I want you to put your finger on the word I have said. Are you ready? Point to chair. Point to light, etc. Now I will place these ear phones on your head and you will again be required to point to the appropriate picture. Be sure and listen to the directions on the tape and if you have any questions, be sure to ask them.

With this tape you will use this score sheet and follow the following instructions: This test is to see how well you can hear words in quiet and in a background of noise. First, here are some practice words. On your answer sheet, underneath the heading "PRACTICE" are five groups (to be called "blocks") of words, six words in each group or block. I will say one of the words in practice block number 1; you are to decide which word it was, and then draw a line through that word. Then I will say a word in practice block number 2 and you will draw a line through the word in that block, and so on. Now here are the practice samples:

nine	1	mine	mix	2	mist	star	3	par	stitch	4	stick	wink	5	wish
line		pine	miss		milk	tar		far	still		stiff	wind		wing
wine		fine	mill		mint	car		char	stilt		sting	with		witch

Now, get ready for the test. The test is divided into parts. Each part contains 50 words. The announcer will say the number of the block, and then say the test word. Just as you did in practice, listen carefully, decide which word in the block was spoken, and draw a heavy line through it. Be sure to do this in each block, even if you are not always certain what the word was. Here is the test. Remember, please guess when necessary.

For those people in Group II, the order of the instructions would be reversed, i.e., the first tape that you will hear will require that you use this score sheet, etc.

APPENDIX C

NAME ______________________ EAR ______________ DATE ______________

MODIFIED RHYME HEARING TEST 2 LIST ______________

1. mop top hop shop cop pop	2. din sin fin pin win tin	3. back bath bass ban bad bat	4. tot lot hot got pot not	5. cut cuff cud cub cuss cup
6. peas peace peach peal peak peat	7. jig big rig pig wig fig	8. safe same save sane sale sake	9. name same game fame came tame	10. dun dud dub dug duck dull
11. sup sud sun sung sub sum	12. tam tang tap tab tan tack	13. law saw raw paw jaw thaw	14. pill pip pin pick pit pig	15. ten hen den pen then men
16. sick pick lick tick wick kick	17. coil oil toil foil soil boil	18. bust dust rust must gust just	19. fold hold cold sold gold told	20. mad mass mat map math man
21. say gay pay way may day	22. shook look book took cook hook	23. fit fill fig fizz fib fin	24. tent sent bent went dent rent	25. pass pat pad pang path pan
26. beat beak beam bean beach bead	27. neat heat beat meat feat seat	28. cave cape came cane case cake	29. seed seethe seek seep seen seem	30. dark hark bark park mark lark
31. feel peel heel eel keel reel	32. pale tale gale male bale sale	33. fang gang bang sang hang rang	34. rave rate raze race ray rake	35. bill will kill hill till fill
36. pane pay pale pave page pace	37. team teak tear teal tease teach	38. red led fed wed bed shed	39. heat heave heath heal hear heap	40. sag sap sass sat sad sack
41. fun nun gun run sun bun	42. pun pup puff putt pub pug	43. lane lace lame late lake lay	44. tip lip sip dip rip hip	45. dill dig din dip dim did
46. vest best test rest nest west	47. fit wit kit sit bit hit	48. kid king kill kit kiss kith	49. sit sin sip sing sill sick	50. but bug bus bun buff buck

APPENDIX D

Percent (Correct) Scores Obtained by the Subjects On the GFW and MRT Tests of Auditory Discrimination

	NOISE SUBTESTS		QUIET SUBTESTS	
	GFW X	MRT Y	GFW X	MRT Y
1.	77	72	100	90
2.	73	68	90	82
3.	57	78	83	86
4.	80	84	93	82
5.	70	76	96	90
6.	90	90	96	98
7.	80	76	90	92
8.	73	80	96	96
9.	77	70	96	84
10.	77	74	90	82
11.	83	82	100	96
12.	73	74	96	92
13.	83	82	100	92
14.	67	76	93	88
15.	86	84	100	94
16.	83	76	96	86
17.	83	74	96	86
18.	86	76	90	92
19.	73	86	100	92
20.	90	82	93	94
21.	73	78	96	84
22.	83	68	100	96
23.	83	74	100	88
24.	76	80	100	90
25.	83	84	93	86
26.	80	80	96	90
27.	80	78	96	84
28.	73	76	96	90
29.	43	58	73	68
30.	86	86	96	96
31.	86	68	96	86
32.	73	82	86	82
33.	80	84	100	94
34.	83	70	93	92
35.	76	80	100	94
36.	83	88	93	88
37.	73	64	90	90
38.	70	72	93	80
39.	63	76	93	80
40.	56	74	73	74