8-1973

A Comparison Of Two Methods Of Speech Reception Threshold Measurement

Richard Allen Oge

This thesis, directed and approved by the candidate's committee, has been accepted by the Graduate Committee of The University of New Mexico in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

A COMPARISON OF TWO METHODS OF SPEECH RECEPTION THRESHOLD MEASUREMENT

Title

Richard Allen Oge

Candidate

Communicative Disorders

Department

Bernard Spolsky

Dean

July 15, 1975

Date

Committee

Richard B. Hood

Chairman

William J. Ryan

Max E. McCallum

A COMPARISON OF TWO METHODS OF SPEECH RECEPTION THRESHOLD MEASUREMENT

BY
RICHARD ALLEN OGE

B.S., Southeast Missouri State University, 1973

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science in
Communicative Disorders
in the Graduate School of
The University of New Mexico
Albuquerque, New Mexico
August, 1975

ACKNOWLEDGMENTS

The writer would like to take this opportunity to thank his committee chairman, Dr. Richard Hood, and the other committee members, Dr. Max McClellan and Dr. William Ryan, for their encouragement and valuable assistance which was necessary for the completion of this study.

R.A.O.

A COMPARISON OF TWO METHODS OF SPEECH RECEPTION THRESHOLD MEASUREMENT

BY
Richard Allen Oge

ABSTRACT OF THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of
Master of Science in
Communicative Disorders
in the Graduate School of

The University of New Mexico

Albuquerque, New Mexico
August, 1975

A COMPARISON OF TWO METHODS OF SPEECH RECEPTION THRESHOLD MEASUREMENT

Richard Allen Oge
Department of Communicative Disorders
The University of New Mexico, 1975

Research has shown that a number of different methods of SRT measurement are now being employed by clinical audiologists. Since different methods are being used to obtain this measurement, variability in test results may be present because of the test administration. This study evaluated two methods of SRT measurement that have been reported to be systematic, valid and reliable. The two methods under investigation are the Chaiklin et al. (1967) 5-dB ascending (because of clinic requirements this study employed a 4-dB interval) and the Wilson et al. (1973) 2-dB descending procedures. The main purpose of this investigation was to present experimental evidence to support the adoption of one or both of these methods for clinical use.

Three investigations have suggested that the CID Auditory Test W-1 be reduced in number of items to achieve greater homogeneity and to increase test precision. This hypothesis, however, had not been clinically tested. A secondary focus of this study was to clinically compare the 27 word list (Curry and Cox, 1966) to the 36 word list (Hirsh et al., 1952) and to report whether the 27 word list yields an SRT statistically different in magnitude or validity than that obtained with the 36 word list.

Speech reception thresholds (SRTs) were obtained on 25 adult males (50 ears) using the Chaiklin 4-dB ascending method and the Wilson 2-dB descending method of measurement. The stimuli were 27 of the 36 spondees dubbed from the CID Auditory Test W-1 records. For the test retest condition, 16 or the original 25 subjects were retested with the above measures, and also with the Chaiklin method paired with the CID 36 word spondee list. The 36 word spondee list was added for the retest procedure in an effort to evaluate whether a statistically different SRT is obtained from using either the 27 word or the 36 word spondee list. The test methodologies, as described by the authors, of the above two methods were carefully followed.

Anaylsis of the test data indicates that there is no statistically significant difference between SRTs obtained with the two methods. Both methods are equal in test retest reliability and validity (i.e., each method shows a high correlation with the two frequency, three frequency and Carhart formulas for pure-tone average). On the average the Chaiklin method is faster to administer than the Wilson method by approximately 27 seconds per SRT. Lastly, there is no statistical difference in the magnitude or validity of SRTs obtained with a 27 word or a 36 word spondee list.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

In an attempt to discover the current methods being utilized in clinical audiology, Martin and Pennington (1971) surveyed a large number of audiologists concerning their administration of the hearing tests available to them. One of the findings of this survey indicates that a number of different procedures are being used to obtain the speech reception threshold (SRT). This fact strongly suggests that the SRT estimate obtained on a given subject may vary between clinics, not because of the subject's hearing sensitivity but because of the audiologist and his procedure of acquiring this measure.

In recent years, several investigators (Chaiklin and Ventry, 1964; Chaiklin et al., 1967; Tillman and Olsen, 1973; Wilson et al., 1973) have recognized this problem and have presented systematic methodologies of SRT measurement as a means to reduce this variability.

The primary purpose of this study is to evaluate two of these proposed methods and to report experimental evidence supporting the adoption of one or both of these methods for clinical use.

Another important aspect of SRT measurement, based on the conclusions of three investigations, was also evaluated in this study. These three investigations (Bowling and

Elpern, 1961; Curry and Cox, 1966; Beattie, et al., 1975) were concerned with the homogeneity of the Central Institute for the Deaf (CID) Auditory Test W-1 spondee words. All three of these studies concluded that some of the 36 words were, in fact, heterogeneous in their intelligibility. That is, the range (10-dB, Bowling and Elpern; 8-dB, Curry and Cox; 7.9-dB, Beattie, et al.) of individual spondee intelligibility was unduly wide and had an adverse effect on the precision of SRT tests. These authors suggested that greater homogeneity would be achieved if the word lists were reduced in number of items. However, their suggestion that a more homogeneous list would yield a different SRT has not been clinically tested.

The secondary focus of this study, then, is to compare a 27 word spondee list (Curry and Cox, 1966) with the standard 36 word spondee list (Hirsh, et al., 1952) to determine if the two lists yield SRTs which differ in magnitude or validity.

CHAPTER II

REVIEW OF THE LITERATURE

Historical Perspective

In 1942, Hughson and Thompson became the first investigators to measure the SRTs of hearing impaired subjects. They demonstrated the agreement between speech and the pure-tone thresholds at 500, 1000, and 2000 Hz.

Between 1942 and 1964, audiologists were obtaining SRTs without the benefit of a systematic methodology for this measure. Jerger et al. (1959) commented on this fact by saying, "the relatively small number of even cursorily described procedures found in the literature were characterized by a certain lack of agreement on some relevant particulars." In an effort to achieve some "objectivity" in their study these authors had to devise their own test methodology.

In 1964, Chaiklin and Ventry compared a descending 2-dB interval to a 5-dB interval method of SRT measurement. At the time of their study most of the current literature (Jerger et al., 1959; Newby, 1958; Tillman and Jerger, 1959) indicated that a 2-dB measurement interval was being used to measure SRTs. Chaiklin and Ventry felt, however, that if a variable is continuously distributed, a small increase (2- or 3-dB) in the measurement interval would have a negligible effect on precision and no effect on reliability. They could,

therefore, see no reason why the SRT measurement interval should be smaller than the 5-dB measurement interval used in obtaining pure-tone thresholds. The purpose of their investigation was two-fold. First, they wanted to report experimental evidence to support the use of 5-dB steps in SRT measurement. Secondly, they wanted to present a valid and reliable test methodology for clinical use.

The results of their study indicated that the 5-dB method of SRT measurement was 1) faster than the 2-dB method, 2) agreed well with the results of the 2-dB method, 3) had a high correlation with the pure-tone average (PTA), and 4) was highly reliable. The authors recommended that the 5-dB method be adopted for clinical use. They also suggested that a 4-dB interval be adopted in clinics that have speech audiometers graduated in 2-dB steps.

One inference was made in the Chaiklin and Ventry (1964) study that was not verified by statistical data. These researchers felt that a 5-dB ascending method of SRT measurement would be quicker to administer than the 5-dB descending method. In a follow-up study, Chaiklin et al. (1967) tested this hypothesis. The results of this investigation did not support the earlier contention that the ascending method was quicker than the descending procedure. The data did, however, demonstrate that the ascending method is a valid and reliable procedure for SRT measurement. Also, the authors pointed out that the "beginning and end points" of the sampling procedure are greatly simplified using the ascending method.

Tillman and Olsen (1973) proposed that a 2-dB

descending method of SRT measurement be adopted for clinical use. The main advantage of this procedure, they say, "is that it confines all clinicians to the same operational definition of threshold, and thus reduces variability in estimates of speech reception threshold produced by variations in this definition." In their article, however, they completely ignore the earlier investigations mentioned above and make the statement, "as will be seen, even the basic issue of the development of a standardized method for establishing the 'threshold for speech' has not been addressed, ..." They support their argument for adoption of their method for clinical use by quoting certain literature and by stating that their method has been "used successfully in clinical settings." However, they do not present any data as to the validity and reliability of the procedure.

A study was presented by Wilson et al. (1973) to substantiate the validity of the method proposed by Tillman and Olsen. These researchers also presented a 5-dB modification of this procedure for use in clinics where the speech audiometers are graduated in 5-dB steps.

The conclusions reached by these investigators were, 1) the 2-dB interval of SRT measurement is just as quick to administer as the 5-dB method, 2) slightly lower thresholds are found using the 2-dB interval, and 3) a high correlation is found to exist between the SRT and the PTA.

Based on the above findings these authors recommended that the Tillman and Olsen method of SRT measurement be adopted for clinical use.

For the purpose of this study the Tillman and Olsen method of SRT measurement will be called the Wilson et al. method. My reasons for this are two-fold. First, the Wilson et al. study presented statistical treatment of the data which the Tillman and Olsen study did not. Secondly, in this study the subject's responses for the 2-dB method of SRT measurement have been recorded on an answer sheet proposed by Wilson et al.

In 1975, Martin and Stauffer modified the Wilson et al. method in order that the SRT could be obtained without prior knowledge of the PTA. Their reason for making this modification was that some situations arise that require determining the SRT without first knowing the PTA. (The ascending method proposed by Chaiklin et al. in 1967 had also eliminated the need for knowing the PTA in order to obtain the SRT.)

Martin and Stauffer compared their modification of the Wilson et al. method to the original procedure and arrived at two conclusions. First, they found that there is no difference between the two methods with respect to the measured SRTs. Secondly, they found no systematic difference between the two methods as to the time required to complete the measure.

Materials

Martin and Pennington (1971) report that the CID Auditory Test W-1 is the most frequently used SRT test material.

This test material was introduced by Hirsh et al. (1952) in an effort to correct some of the deficiencies that had been found in the Harvard Psycho-Acoustic Laboratory (PAL) word lists. Hirsh and his associates reduced the PAL 84

word list to a total of 36 homogeneous spondees (two syllable words with equal stress on each syllable). Using these 36 words they made the CID Auditory Tests W-1 and W-2 word lists. Six randomizations of each list were dubbed onto magnetic tape. The W-1 word lists were then presented to six inexperienced and six experienced listeners. After evaluating the test results from these 12 listeners the authors reported:

> The articulation score rose from 0 to 100 percent within a range of about 20 dB. There is an increase from 20 to 80 percent within a range of 8 dB and throughout this range the slope or rate of rise in score is about 8 percent per dB. Since the threshold falls on the steepest part of the function, it is crossed very abruptly and, therefore, can be very sensitively determined with this test. (Hirsh et al., 1952 p. 324)

Hirsh et al. concluded that this test material seemed well suited for the measurement of SRTs.

Bowling and Elpern (1961), however, noticed that certain items on the CID test tended to be consistently more or less intelligible than others and caused an adverse effect on the precision of the test results. In an effort to verify their hypothesis, they devised a study to estimate the range of the intelligibility of the 36 words contained in the CID W-1 List. Their data indicated that a greater homogeneity and more test precision would be gained if the list was reduced to 22 words. However, a number of conditions existed that may have influenced their test results. First, the subjects were not familiarized with the test material. The study by Tillman and Jerger (1959) had shown that subjects

given prior knowledge of the test material would yield spondee thresholds 4-5-dB lower than subjects not given this knowledge. Secondly, the subjects were not encouraged to guess at the words. The third factor which may have influenced their results is that a small number of subjects (24) were used in the experiment.

Curry and Cox (1966) also investigated the range of intelligibility of the CID W-1 List. Their study had some important modifications that the Bowling and Elpern study did not contain. They used 50 subjects, familiarized them with the test material, and encouraged guessing. The results obtained by these researchers agreed with the earlier conclusions of Bowling and Elpern with the exception that they recommended reducing the list to 27 words rather than to 22 words.

Both of the preceding studies employed magnetic tape dubs of the CID Auditory Test W-1. In 1975, Beattie et al. investigated the intelligibility of the CID spondees presented via monitored live voice. Three experienced examiners each tested a group of 25 subjects. These 75 subjects were familiarized with the test words and were encouraged to guess if they were not sure of the word. The results obtained from this study agreed with the research of both Bowling and Elpern and Curry and Cox with the exception that Beattie et al. recommended reducing the list to 18 words to achieve a greater homogeneity of the list.

Relationships Between the Speech Reception Threshold and the Pure-tone Average

Many researchers have verified the interdepedence between the threshold for speech and the sensitivity of the mid-frequency pure-tones since it was first reported by Hughson and Thompson (1942). It is agreed by most audiologists that this phenomenon exists; the major area of disagreement concerns the choice of the weighting system which should be used to predict the subject's threshold for spondees. As the literature on this topic is too voluminous to cover in detail, only three of the more popular methods of figuring the PTA will be reviewed.

Carhart (1946) found that SRTs obtained with the PAL spondees and with connected speech were basically the same. After further research, he concluded that the simple method of averaging the thresholds for 512, 1024, and 2048 Hz was an accurate method of calculating the subject's threshold for spondees.

Another method of calculating the spondee threshold from the pure-tone average was devised by Fletcher (1950). He recommended that the best two frequencies between 500 and 2000 Hz be averaged to obtain the SRT of the subject.

Carhart (1971) examined a large number of clinical records in which the pure-tone thresholds and speech reception thresholds had been obtained on audiometers calibrated to the American National Standards Institute (ANSI) 1969 norms. He reported that when the audiometric pattern is not taken into account, a good formula for predicting SRTs is to average the

thresholds obtained at 500 and 1000 Hz and to subtract the correction constant of 2-dB from this average.

Summary

A number of different methods of SRT measurement are now being employed by clinical audiologists. Since different methods are being used to obtain this measurement, variability in test results may be present because of the test administration. This study proposes to evaluate two methods of SRT measurement that have been reported to be systematic, valid and reliable. The two methods under investigation are the Chaiklin et al. (1967) 5-dB ascending (because of clinic requirements this study employed a 4-dB interval) and Wilson et al. (1973) 2-dB descending procedures. The main purpose of this investigation is to present experimental evidence to support the adoption of one or both of these methods for clinical use.

Three investigations have suggested that the CID Auditory Test W-1 be reduced in number of items to achieve greater homogeneity and to increase test precision. This hypothesis, however, has not been clinically tested. A secondary focus of this study is to clinically compare the 27 word list (Curry and Cox, 1966) to the 36 word list (Hirsh et al., 1952) and to report whether the 27 word list yields an SRT statistically different in magnitude or validity than that obtained with the 36 word list.

Statement of the Problem

The main objective of this study is to compare two existing methods of SRT measurement and to evaluate their respective correlations with three methods of PTA.

A secondary objective is to clinically compare the 27 word spondee list with the 36 word spondee list to determine if either one of these lists produces an SRT that correlates better with the PTA.

More specifically the following questions were asked:

1. Which method yields better results as to SRT-PTA agreement, the Chaiklin et al. (1967) 4-dB ascending method of SRT measurement or the Wilson et al. (1973) 2-dB descending method of SRT measurement?
2. Which method is quicker in establishing the SRT measurement, the Chaiklin method or the Wilson method?
3. Do SRTs established with the 27 word W-1 spondee list yield a correlation with the PTA that differs significantly from the correlation established with the 36 word W-1 spondee list?
4. Do SRTs established with the 27 word W-1 spondee list yield results that are statistically different as to magnitude or validity than the results obtained with the 36 word W-1 spondee list?

CHAPTER III

PROCEDURE

Subjects

The subjects for this study were adult male patients which were seen at the Albuquerque Veterans Administration Hospital, Speech Pathology and Audiology Clinic. The data are based on 50 ears from 25 subjects who were incorporated into the study as they appeared for audiological services. These services consisted of routine audiometric examinations, hearing aid evaluations, compensation ratings, and special diagnostic testing. The subjects ranged in age from 26 to 59 years with a mean age of 42.7 years. The types of hearing loss for the subjects are shown in Table 1. It is important to note, however, that of the 50 ears tested, 32 had PTAs which were 26-dB or lower. Also, 18 of these 32 ears showed a threshold of 15-dB or greater in at least one of the speech range frequencies.

The following criteria were used to exclude an ear from being used in the study:

1. Any ear, nose, or throat condition that might cause abnormal threshold fluctuations;
2. A speech discrimination score lower than 60%;
3. Subjects over 62 years of age;
4. Subjects exhibiting a non-organic hearing loss.

TABLE I

TYPES OF HEARING LOSS

Category		N
Normal across all frequencies		7
Mixed		4
Normal in speech range	3	
15-dB or greater loss in speech range	1	
Sensori-neural		39
Normal in speech range	22	
15-dB or greater loss in speech range	17	

Instrumentation and Calibration

Data were collected in two double-walled Industrial Acoustic Company sound treated suites. Two Grason-Stadler Model 1701 audiometers equipped with Telephonics TDH-49 earphones in MX-41/AR cushions were used for all threshold measurements. Two, Sony Model TC-366, two-channel tape decks were used as speech signal inputs. The calibration of the speech and pure-tone circuits was checked before, during, and immediately after completion of the study with a Bruel and Kjaer, Model 158, audiometer calibrator. Because the TDH-49 earphones were used the speech and pure-tone circuits were calibrated to the norms set forth by Delany and Whittle (Grason-Stadler Co., 1971 p. 26). A comparison between the above norms and the ANSI 1969 norms can be found in Table 2.

Pure-tone Threshold Measurement

Pure-tone thresholds were measured with an ascending technique. Initial presentation of signals was begun at -10dB HTL. Moving in an ascending manner one stimulus was presented at each succeeding 10-dB increment until a response was elicited from the subject. The hearing level dial was then attenuated 10-dB. One stimulus was then presented at each 5-dB increment until another response was elicited. This pattern was then continued (descending 10-dB after a response and ascending in 5-dB steps with one stimulus at each step until another response was made) until three stimuli had been responed to at one level. The three responses did not have to be in

TABLE 2

A COMPARISON OF THE DELANY AND WHITTLE AND THE ANSI 1969 AUDIOMETER CALIBRATION NORMS

	Frequency in Hertz									
Norm	250	500	1000	1500	2000	3000	4000	6000	8000	SPEECH
Delany and Whittle	26.5	13.5	7.5	7.5	11.0	9.2	10.0	13.0	14.5	20.0
ANSI 1969	25.5	11.5	7.0	6.5	9.0	10.0	9.5	15.5	13.0	20.0

succession. The level at which the three responses were made was considered to be threshold for that frequency.

Stimuli

Speech signals were magnetic tape dubs of the Technisonic Studio's recordings of CID Auditory Test W-1, lists A through D (Hirsh et al., 1952). For initial SRT testing, a 27 word list (Appendix A), as recommended by Curry and Cox (1966) was used. The 27 word list was a tape dub of the W-1 recording with the nine least homogeneous words eliminated. The test retest data were collected on 16 ears using the 27 word list. The 36 word spondee list was also used at this time as a means of comparing the results obtained with this list to the results obtained with the 27 word list.

Test Procedure

First, pure-tone air and bone conduction thresholds were obtained. Then automatic Bekesy tracings were obtained to rule out subjects with a possible non-organic hearing loss. Next, the subject was read the instructions found in Appendix B which are appropriate for both methods of SRT measurement under study. The subject then read aloud the 36 spondee words which were typed on 4" x 6" cards. If the subject could not read the words, they were read to him, and he repeated them back to the tester. This procedure familiarized the subject with the words to be used and reduced learning effects (Tillman and Jerger, 1959). After the subject was familiarized with the spondees, the speech stimuli were calibrated by set-

ting the VU meter to zero (0) on the final word of the phrase "Say the word." Speech reception thresholds were then measured using the Wilson et al. (1973) 2-dB descending and the Chaiklin et al. (1967) 4-dB ascending methods. For these two methods the 27 word spondee list (Curry and Cox, 1966) was used. Speech discrimination scores were obtained as the final part of the initial evaluation.

To control for order effects, the SRT methods were randomly assigned (Appendix C) to each subject. Also, all procedures were performed by the author to insure uniformity in the testing procedures.

Since the subjects could not be recalled, the retest procedures were administered after all initial testing was completed and after the subject had a brief rest period.

The retest procedure was performed on one ear of 16 randomly selected subjects. The procedures used were the same as the SRT measurements listed above with the addition of the 36 word spondee list paired with the Chaiklin et al. ascending technique. The retest SRT methods were also randomly assigned to the retest ears (Appendix D).

All subject responses were recorded on an answer sheet, which incorporated the scoring methods used by both of the original investigators in their respective studies (Appendix E).

Procedure for 4-dB Ascending SRT Measurement

The hearing level dial was set to -10dB HTL and the first stimulus was presented. Succeeding stimuli were presented

in ascending 10-dB steps until the subject made a correct response. The hearing level dial was then reduced 16-dB (personal communication to M. McClellan from J. Chaiklin, 1968) below the correct response, and threshold sampling was begun in 4-dB ascending steps. Sampling was continued at each level until the 50% criterion (3 out of 6 correct responses) was achieved. For example, if four words were missed at a given level, the 50% criterion would not be met, and sampling would be discontinued at that level and continued at the next higher level (Chaiklin et al., 1967).

Procedure for 2-dB Descending SRT Measurement

The hearing threshold level was set to 30-dB HTL above the estimated threshold of the subject, and the first stimulus was presented. If the subject gave a correct response, the hearing threshold level was lowered 10-dB and another stimulus was presented. This procedure continued until the subject gave an incorrect response. Then, at the same intensity level a second stimulus was presented. If that stimulus was also missed, the intensity level was raised 10-dB, and the test was begun. If, however, the second stimulus was repeated correctly, the intensity level was lowered 10-dB, and the above process was repeated until two stimuli were missed at the same intensity level. The threshold testing started at 10-dB above the intensity level at which two stimuli were missed. This starting level had to be high enough, however, to allow the subject to correctly respond to at least five of the first six stimuli presented. If this criterion was not met, the intensity

level was raised 6-dB above the initial starting level. Once the starting level had been determined, two successive stimuli were presented to the subject at that intensity level. The intensity level was then lowered 2-dB and two more stimuli were presented. This process continued at 2-dB levels of attenuation until the subject missed at least five of the last six stimuli presented to him. To determine the subject's SRT, the number of correct responses were counted, decreased by one (1), and this number was subtracted from the hearing level setting at which the test was begun (Tillman and Olsen, 1973).

Test Retest

Sixteen of the original 50 ears were chosen randomly for the test retest evaluation. The test retest procedure followed the same pattern as the initial SRT measurements. The two methods under study were randomly assigned to the 16 retest ears. A third procedure, the Chaiklin et al. 4-dB ascending method, paired with the 36 word spondee list was also randomly assigned to the 16 retest ears. By adding this third method of SRT measurement, the effect that the 27 word as compared to the 36 word spondee list has on the correlation between the SRT and the PTA could be evaluated.

Pure-tone Averages

The SRT scores obtained with the two methods of measurement under study were compared to three methods of calculating the PTA. These three methods were:

1. Average of the pure-tone thresholds at 500, 1000, and 2000 Hz (three frequency average);
2. Average of the best two pure-tone thresholds at 500, 1000, and 2000 Hz (two frequency average);
3. Average of the pure-tone thresholds at 500 and 1000 Hz minus 2-dB (Carhart Method).

Measurement Time

The measurement time of the two SRT methods was evaluated by counting the number of stimuli required to establish the SRT for each ear. This figure can be converted into time by multiplying the number of stimuli needed to establish the SRT by six seconds (the time interval for the carrier phrase, spondee, and the silence before the next stimulus is presented).

Data Reduction and Analysis

Measures of central tendency (mean) and variability (standard deviation) were computed. These data were then submitted to analysis by t-test. The Pearson product-moment correlation coefficient was used to determine correlations between the:

1. SRT methods and the three methods of PTA;
2. The results obtained with the 2-dB and 4-dB SRT methods;
3. Test retest reliability

The standard error of the estimate was used to determine how well the SRT predicts the PTA.

CHAPTER IV

RESULTS

Comparison of Measurement Methods

The following results were found when the Chaiklin et al. (1967) 4-dB ascending method and the Wilson et al. (1973) 2-dB descending method of SRT measurement test data were compared.

The mean SRT obtained using the Wilson method was 13.24dB and the mean SRT for the Chaiklin method was 12.36dB. This difference of .88-dB between the two means was not statistically significant (t=.05). Very similar standard deviations were found for the two methods. The correlations of the SRTs obtained using the two methods was very high (r=.90). See Table 3 for a summary of these details.

TABLE 3

A COMPARISON OF SRTs OBTAINED USING THE WILSON AND CHAIKLIN METHODS (N=50)

Method	Mean	SD	t	r
Wilson	13.24	8.19	.537*	.90
Chaiklin	12.36	7.99		

* Not significant at the .05 level

Measurement Method versus Pure-tone Average

Table 4 shows a summary of the correlations between the SRTs obtained with the two methods of measurement and the two frequency, three frequency, and the Carhart method of PTAs. The mean PTAs for the three averages are also shown.

As can be seen from the results in Table 4, the two methods of SRT measurement produce essentially the same high degree of correlation with the three methods of PTAs, although the three frequency formula predicted the SRTs slightly less accurately than the two frequency formulas. The values of the standard error of the estimate also point up this small difference in predictability.

The important findings are that both methods yield equivalent mean SRTs, standard deviations, and equally valid and predictable SRTs when a particular formula for pure-tone average is used as the validating criterion.

Measurement Time

The mean number of stimuli required to obtain the SRT with the Wilson method was 26.0 and the mean number of stimuli required for the Chaiklin method was 21.4. This difference (4.6 stimuli) is equivalent to a mean time difference of approximately 27 seconds.

Analysis of the data shows that three SRTs were obtained using the same number of stimuli for both methods of measurement. Fifteen SRTs were obtained in which the Chaiklin method took longer (range = 1 to 12 more stimuli required)

TABLE 4

MEAN PTAs (in dB), AND CORRELATIONS
BETWEEN SRTs AND PTAs (N=50)

Method	2-Freq. PTA			3-Freq. PTA			Carhart PTA		
	mean	r	SE est.	mean	r	SE est.	mean	r	SE est.
Wilson	10.98	.91	3.40	15.88	.83	4.57	10.00	.91	3.40
Chaiklin		.96	2.23		.82	4.58		.89	3.64

while 32 SRTs were obtained with the Wilson method requiring more stimuli (range = 2 to 25 more stimuli required).

Test Retest

Sixteen ears were retested for this part of the study. The retest SRTs were all within 4-dB of the original SRTs using the Chaiklin method. With the Wilson method, 12 were within 4-dB, three were within 8-dB, and one was more than 8-dB. The correlation between the initial test and the retest was high for both methods (Chaiklin, r=.94; Wilson, r=.87). Although these findings on a small number of subjects slightly favor the Chaiklin method, the main finding is that the test retest reliability is very high for both methods.

36 versus 27 Word Lists

A third SRT measurement was added to the procedure during the test retest evaluation: the Chaiklin et al. (1967) 4-dB ascending technique using the full 36 word spondee list. This measurement was added to evaluate whether SRTs obtained with the 27 word list were statistically different from SRTs obtained with the 36 word list. Table 5 shows a summary of the results that were obtained.

The differences between the means of these three methods were not statistically significant. Table 6 shows a summary of the t-test comparisons.

A high correlation was found to exist between the SRTs obtained with the three methods of measurement and the

TABLE 5

A COMPARISON OF THE CHAIKLIN, 27 WORD ASCENDING METHOD, WILSON, 27 WORD DESCENDING METHOD AND THE CHAIKLIN, 36 WORD ASCENDING METHOD (N=16)

	Method		
	Chaiklin 4-dB 27 word	Wilson 2-dB 27 word	Chaiklin 4-dB 36 word
Mean	11.50	11.06	9.75
SD	7.22	6.45	6.85

TABLE 6

A COMPARISON OF t-TEST RESULTS FOR THE CHAIKLIN, 27 WORD ASCENDING METHOD, WILSON, 27 WORD DESCENDING METHOD, AND THE CHAIKLIN, 36 WORD ASCENDING METHOD (N=16)

Methods	t	p
Chaiklin 27 word versus Wilson 27 word	.174	NS
Wilson 27 word versus Chaiklin 36 word	.539	NS
Chaiklin 27 word versus Chaiklin 36 word	.680	NS

PTAs. A summary of these results are found in Table 7.

TABLE 7

MEAN PTAs (In dB), AND CORRELATIONS BETWEEN PTAs AND RETEST SRTs (N=16)

	2-Freq. PTA		3-Freq. PTA		Carhart PTA	
Method	mean	r	mean	r	mean	r
	9.93		15.93		8.87	
Chaiklin 27 word		.91		.93		.93
Wilson 27 word		.90		.82		.90
Chaiklin 36 word		.93		.90		.93

One conclusion that can be reached from the data shown in Tables 5, 6, and 7 is that, under clinical conditions, there is not a significant difference between SRTs obtained using the 27 word spondee list and SRTs obtained using the 36 word spondee list. Another conclusion is that the methods yielded SRTs which were highly correlated with the PTAs.

CHAPTER V

DISCUSSION

Briefly, the results of the present study may be summarized as follows:

1. Both the Chaiklin et al. (1967) 4-dB ascending SRT method and the Wilson et al. (1973) 2-dB descending SRT method of measurement yield results that are statistically indistinguishable.
2. The Chaiklin method of SRT measurement is faster to administer than the Wilson method, on the average.
3. The test retest reliability of both methods is very high.
4. Both the Chaiklin method and the Wilson method show high correlations with the two frequency, three frequency, and Carhart formulas of PTA.
5. There is no statistical difference between SRTs obtained with the 27 word spondee list or with the 36 word spondee list.

On the basis of the above findings, it would seem that either method of SRT measurement could be adopted by a clinician without the loss of either validity or reliability. If the clinic, however, is one in which the time of test administration is an important factor, I would recommend that the Chaiklin et al. (1967) 4-dB ascending method be used.

By using this method approximately one minute per patient would be saved.

The three studies (Bowling and Elpern, 1961, Curry and Cox, 1966; Beattie et al., 1975) on spondee word intelligibility concluded that the CID Auditory List W-1 should be reduced in number of items to achieve a greater homogeneity of the list. Although the list may be heterogeneous in respect to word intelligibility, this fact does not seem to affect the precision of the test. The data from this study indicate that there is no statistical difference between SRTs arrived at by using the 27 word spondee list or the 36 word spondee list.

Although no statistical differences were found concerning the validity and reliability of the two methods of SRT measurement under investigation, it is possible that the characteristics of the sample influenced the results. Specifically, this study was confined to using subjects from the Veterans Administration Hospital population. The very nature of the hearing losses in this population could have had an influence on the results of the study. Most of the subjects had high frequency, sensori-neural type hearing losses with nearly normal hearing through the speech range. This homogeneity of audiometric configurations undoubtedly accounts for the fairly low SRT scores and may partially account for the good agreement of their scores with the two frequency and Carhart method of PTAs.

Some subjective observations were also noticed by the author during the course of this investigation. The Chaiklin et al. (1967) methods of SRT measurement seemed

easier to administer than the Wilson et al. (1973) method for the following reasons. First, the starting and ending points of the Chaiklin method are always easy to determine, whereas the starting point for the Wilson method is often hard to establish. Secondly, a form must be used to record the responses with the Wilson method, whereas it is not necessary with the Chaiklin method (although Chaiklin et al. recommend one for convenience). Lastly, the Wilson method requires that calculations be made to obtain the SRT, while the SRT using the Chaiklin method is the first level at which the criterion is met.

I would also recommend that the starting level of the Wilson method be raised from 10-dB to 16-dB above the level at which two consecutive stimuli are missed. During this study the starting level had to be raised higher than the initial 10-dB for 66% of the original 50 ears and for 75% of the 16 retest ears. The reason for this was that the criterion of correctly identifying five of the first six stimuli could not be met.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Summary

Speech reception thresholds (SRTs) were obtained on 25 adult males (50 ears) using the Chaiklin 4-dB ascending method and the Wilson 2-dB descending method of measurement. The stimuli were 27 of the 36 spondees dubbed from the CID Auditory Test W-1 records. For the test retest condition, 16 of the original 25 subjects were retested with the above measures, and also with the Chaiklin method paired with the CID 36 word spondee list. The 36 word spondee list was added for the retest procedure in an effort to evaluate whether a statistically different SRT is obtained from using either the 27 word or the 36 word spondee list. The test methodologies, as described by the authors, of the above two methods were carefully followed.

Analysis of the test data indicates that there is no statistically significant difference between SRTs obtained with the two methods. Both methods are equal in test retest reliability and validity (i.e., each method shows a high correlation with the two frequency, three frequency and Carhart formulas for pure-tone average). On the average the Chaiklin method is faster to administer than the Wilson method by approximately 27 seconds per SRT. Lastly, there is no statistical

difference in the magnitude or validity of SRTs obtained with a 27 word or a 36 word spondee list.

Conclusions

Based on the results of this investigation, the author can recommend both the Chaiklin et al. (1967) 4-dB ascending method and the Wilson et al. (1973) 2-dB descending method of SRT measurement for clinical use. Both of these methods are highly valid and reliable. If, however, time is an important factor in the test administration it is recommended that the Chaiklin method be used. This method will save the tester approximately one minute per subject, on the average. Under clinical conditions either the 27 word spondee list or the 36 word spondee list may be used since the results obtained with these two lists show no statistical difference.

The Chaiklin method is easier to administer, in my opinion, for the following reasons:

1. The starting and ending points are easy to determine.
2. A response record form is not a necessity.
3. Arithmetic calculations are not required.

Suggestions for Future Research

In view of the present findings, several suggestions for future research may be made:

1. Replication of the present study with a subject population that exhibits audiometric configurations that are more heterogeneous.

2. An investigation to determine if either the ascending or the descending method of SRT measurement identifies non-organic hearing losses more accurately.
3. An investigation which compares the validity of SRTs obtained using the full 36 word CID list to those obtained using abbreviated lists which allegedly contain words that are more homogeneous in intelligibility.

APPENDICES

APPENDIX A

CURRY AND COX
27 WORD SPONDEE LIST

ICEBERG

AIRPLANE

ARMCHAIR

PLAYGROUND

DRAWBRIDGE

WOODWORK

HARDWARE

COWBOY

BIRTHDAY

GREYHOUND

EARDRUM

SUNSET

NORTHWEST

STAIRWAY

SIDEWALK

RAILROAD

DAYBREAK

DOORMAT

SCHOOLBOY

INKWELL

OATMEAL

MUSHROOM

MOUSETRAP

WHITEWASH

FAREWELL

PADLOCK

TOOTHBRUSH

APPENDIX B

SUBJECTS' TEST INSTRUCTIONS

The purpose of this test is to find the faintest level at which you hear words. You are going to hear some two syllable words such as "baseball" and "mousetrap." Your job is to repeat each two syllable word no matter how faint the word may be. For example, if you hear, "say the word 'baseball,'" you should just repeat "baseball." The words may be very faint, but please continue repeating them as well as you can until I tell you the test is over. It's important that you guess at the word, even if you're not sure of it. Do you have any questions?

Now I want you to take this stack of cards and read the words aloud to me. These are the words that you'll hear during the test. I want you to become familiar with them. (AFTER READING CARDS)

Remember, it's very important that you guess, no matter how faint a word may be or even if you only hear part of the word.

Chaiklin and Ventry, (1964) p. 55.

APPENDIX C

RANDOMIZATION OF INITIAL TEST
(* indicates test retest ear)

Number	1st Test	Spondee List	2nd Test	Spondee List
* 1	C	A	W	C
2	W	B	C	D
3	W	D	C	C
4	C	D	W	B
5	C	C	W	B
6	W	B	C	A
* 7	W	D	C	C
8	C	B	W	D
* 9	C	C	W	B
* 10	W	B	C	A
11	C	C	W	D
12	C	C	W	B
13	W	B	C	A
14	W	B	C	A
* 15	W	D	C	A
16	W	D	C	A
17	C	C	W	B
18	C	A	W	D
19	W	B	C	A
20	C	B	W	A
21	C	B	W	D
22	C	C	W	B
23	W	A	C	C
24	W	A	C	B
25	W	B	C	A
26	C	D	W	A
* 27	C	B	W	D
28	C	C	W	B
29	C	D	W	A
30	W	A	C	C
31	C	D	W	B
* 32	W	B	C	C
33	W	C	C	B
* 34	C	A	W	B
35	W	B	C	D
* 36	C	A	W	B
37	W	B	C	D
38	C	D	W	A
* 39	C	B	W	A
40	C	C	W	D
* 41	C	A	W	B
42	W	D	C	B
43	C	A	W	B
44	C	A	W	B
45	C	B	W	A

APPENDIX C
(Continued)

Number	1st Test	Spondee List	2nd Test	Spondee List
46	W	C	C	B
* 47	C	C	W	D
48	W	D	C	B
49	C	D	W	A
* 50	C	C	W	A

W is Wilson et al. (1973) method with 27 word spondee list.
C is Chaiklin et al. (1967) method with 27 word spondee list.

APPENDIX D

RANDOMIZATION OF TEST RETEST

Number	1st Test	Spondee List	2nd Test	Spondee List	3rd Test	Spondee List
1	W	D	C	B	C*	A
2	C	D	W	C	C*	B
3	W	C	C	B	C*	D
4	W	A	C	B	C*	C
5	W	A	C	C	C*	D
6	C*	C	C	B	W	A
7	C	B	C*	C	W	D
8	W	C	C	D	C*	A
9	C*	B	C	C	W	A
10	W	D	C*	B	C	C
11	C	D	W	C	C*	B
12	C*	D	W	A	C	B
13	C*	D	W	B	C	C
14	W	C	C*	A	C	D
15	C	B	C*	D	W	C
16	C	B	C*	D	W	A

W is Wilson et al. (1973) method with 27 word spondee list.
C is Chaiklin et al. (1967) method with 27 word spondee list.
C* is Chaiklin et al. (1967) method with 36 word spondee list.

APPENDIX E

SCORE SHEET

Audiometer 1 2 No.____

NAME:_____ DATE:_____ AGE:_____

EAR: R L DISCRIMINATION: R____% L____%

PTA:

	2 Freq.	3 Freq.	Carhart
R			
L			

HTL	RIGHT	ERROR 1	ERROR 2	ERROR 3	ERROR 4

SRT:_____ No. Stim:_____

INT.			INT.		
0			0		
8			8		
6			6		
4			4		
2			2		
0			0		
8			8		
6			6		
4			4		
2			2		

X = Incorrect
✓ = correct

SRT:_____ No. Stim:_____

APPENDIX F

INITIAL TEST RESULTS

Subject	Age	SRT		PTA		Carhart	Number of Stimuli	
		Chaiklin	Wilson	2-Freq.	3-Freq.		Chaiklin	Wilson
1	56	18	20	15	17	13	19	19
2	56	30	20	25	25	23	30	18
3	44	12	11	8	17	6	16	36
4	44	6	10	10	25	8	21	21
5	26	20	20	15	20	13	26	23
6	26	20	22	20	28	18	25	20
7	27	2	3	-3	3	-5	15	22
8	27	10	7	3	12	1	26	27
9	41	22	23	25	40	23	20	34
10	41	22	22	28	37	25	18	24
11	35	12	12	5	7	6	18	16
12	35	8	11	8	8	6	20	27
13	43	4	6	5	7	8	11	20
14	43	8	7	5	12	3	17	29
15	37	8	3	3	3	3	14	28
16	37	4	6	5	5	3	28	30
17	28	8	11	10	10	8	14	23
18	28	10	13	10	12	13	25	20
19	45	10	7	5	5	3	26	22
20	45	10	8	5	7	3	26	27
21	42	6	5	5	7	3	22	25
22	42	6	7	10	12	8	20	34
23	30	26	26	23	18	21	23	33
24	30	26	31	23	27	25	24	20
25	57	6	9	5	5	3	22	23

APPENDIX F
(Continued)

Subject	Age	SRT		PTA		Carhart	Number of Stimuli	
		Chaiklin	Wilson	2-Freq.	3-Freq.		Chaiklin	Wilson
26	57	6	6	5	10	3	20	36
27	53	16	14	10	22	8	26	41
28	53	20	27	23	32	21	32	27
29	46	26	21	25	29	26	23	31
30	46	18	16	23	23	23	15	40
31	26	16	17	10	23	8	24	22
32	26	14	20	10	32	16	16	29
33	46	6	7	5	5	3	25	28
34	46	10	11	10	12	8	27	35
35	53	2	8	5	7	6	16	20
36	53	6	8	3	5	3	22	22
37	58	16	11	13	27	11	23	21
38	58	38	43	33	40	31	19	31
39	48	8	17	10	12	11	15	24
40	48	6	12	8	8	8	22	21
41	40	6	5	3	3	1	22	16
42	40	8	6	5	8	3	34	28
43	45	6	6	3	8	1	21	28
44	45	6	9	3	10	1	23	26
45	39	10	10	8	8	6	24	23
46	39	14	18	15	23	13	15	28
47	45	2	6	3	8	6	18	17
48	45	6	8	3	8	6	21	28
49	59	16	13	15	25	13	25	33
50	59	22	28	25	37	23	18	24

APPENDIX G

TEST RETEST RESULTS

Subject	SRT C-27	SRT W-27	SRT C-36	PTA 2-Freq.	PTA 3-Freq.	Carhart	Number of Stimuli C-27	Number of Stimuli W-27	Number of Stimuli C-36
1	16	15	18	15	17	13	26	16	17
2	2	2	2	-3	3	-5	18	22	20
3	22	20	18	25	40	23	20	24	14
4	22	20	20	28	37	25	18	21	29
5	2	6	2	5	5	3	16	24	18
6	10	8	6	5	7	3	26	22	21
7	6	10	2	5	5	3	21	22	16
8	14	10	12	10	22	8	30	33	19
9	18	11	16	12	32	16	13	33	23
10	6	7	6	10	12	8	22	30	23
11	10	9	6	3	5	3	26	29	21
12	12	9	12	10	12	11	30	26	19
13	6	8	6	3	3	1	22	24	21
14	6	4	2	3	10	1	18	17	19
15	6	10	6	3	8	6	22	24	23
16	26	28	22	25	37	23	26	24	19

C-27 is the Chiklin method with the 27 word spondee list.
W-27 is the Wilson method with the 27 word spondee list.
C-36 is the Chaiklin method with the 36 word spondee list.

BIBLIOGRAPHY

Beattie, R.C., Svihovec, D.V., and Edgerton, B.J., Relative intelligibility of the CID spondees as presented via monitored live voice. J. Speech Hearing Dis., 40, 84-91 (1975).

Bowling, L.S., and Elpern, B.S., Relative intelligibility of items on CID auditory test W-1. J. Aud. Res., 1, 152-157 (1961).

Carhart, R., Monitored live-voice as a test for auditory acuity. J. Acoust. Soc. Amer., 17, 339-341 (1946).

Carhart, R., Observations on relations between thresholds for pure tones and for speech. J. Speech Hearing Dis., 36, 476-483 (1971).

Carhart, R., and Jerger, J.F., Preferred method of clinical determination of pure-tone thresholds. J. Speech Hearing Dis., 24, 330-345 (1959).

Chaiklin, J.B., Personal communication to M. McClellan, (1968).

Chaiklin, J.B., Font, J., and Dixon, R.F., Spondee thresholds measured in ascending 5-dB steps. J. Speech Hearing Res., 10, 141-145 (1967).

Chaiklin, J.B., and Ventry, I.M., Spondee threshold measurement: A comparison of 2- and 5-dB methods. J. Speech Hearing Dis. 29, 47-59 (1964).

Curry, E.T., and Cox, B.P., The relative intelligibility of spondees. J. Aud. Res., 6, 419-424 (1966).

Fletcher, H., A method of calculating hearing loss for speech from the audiogram. Acta. Otolaryngol. Suppl., 90, 26-37 (1950).

Grason-Stadler Co., 1701 Audiometer Instruction Manual. Concord, Mass. (1971).

Hirsh, I.J., Davis, H., Silverman, S.R., Reynolds, E.G., Eldert, E., and Benson, R.W., Development of materials for speech audiometry. J. Speech Hearing Dis., 17, 321-337 (1952).

Hughson, W. and Thompson, E.A., Correlation of hearing acuity for speech with discrete frequency audiograms. Arch. Otolaryngol., 36, 526-540 (1942).

Jerger, J.F., Carhart, R., Tillman, T.W., and Peterson, J.L., Some relations between normal hearing for pure tones and for speech. J. Speech Hearing Res., 2, 126-140 (1959).

Martin, F.N., and Pennington, C.D., Current trends in audiometric practices. ASHA., 13, 671-677 (1971).

Martin, F.N., and Stauffer, M.L., A modification of the Tillman-Olsen method for obtaining the speech reception threshold. J. Speech Hearing Dis., 40, 25-28 (1975).

Newby, H.A., Audiology: Principles and Practice. New York: Appleton-Century-Crofts, (1958).

Tillman, T.W., and Jerger, J.F., Some factors affecting the spondee threshold in normal hearing subjects. J. Speech Hearing Res., 2, 141-146 (1959).

Tillman, T.W., and Olsen, W.O., Speech audiometry. Ch. 2 In J.F. Jerger (Ed.), Modern Developments in Audiology (2nd edition) New York: Academic Press, (1973).

Wilson, R.H., Morgan, D.E., and Dirks, D., A proposed SRT procedure and its statistical precedent. J. Speech Hearing Dis., 38, 184-191 (1973).