Development and Standardization of Spondee and Phonetically Balanced (PB) Word Lists in Mizo Language

Lalh Mangaiahi

Register No: 07AUD007

A Dissertation Submitted in Part Fulfillment of

Final year M.Sc (Audiology),

University of Mysore, Mysore.

ALL INDIA INSTITUTE OF SPEECH AND HEARING

NAIMISHAM CAMPUS, MANASAGANGOTHRI

MYSORE-570006

May 2009

Certificate

This is to certify that this dissertation entitled **"Development and Standardization of Spondee and Phonetically Balanced (PB) Word Lists in Mizo Language"** is a bonafide work in part fulfillment for the degree of master of (Audiology) of the student (Registration No. 07AUD007). This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore

May, 2009

Dr.Vijayalakshmi Basavaraj

Director

All India Institute of Speech and Hearing

Manasagangothri, Mysore -570006.

Certificate

This is to certify that this dissertation entitled **"Development and Standardization of Spondee and Phonetically Balanced (PB) Word Lists in Mizo Language"** is a bonafide work in part fulfillment for the degree of master of (Audiology) of the student (Registration No. 07AUD007). This has been carried out under my guidance and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore

May, 2009

Dr. Vijayalakshmi Basavaraj

Director & Guide

All India Institute of Speech and Hearing

Manasagangothri, Mysore -570006.

Declaration

This Dissertation entitled **"Development and Standardization of Spondee and Phonetically Balanced (PB) Word Lists in Mizo Language"** is the result of my own study under the guidance of Dr. Vijayalakshmi Basavaraj, Director, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier in any other University for the award of any Diploma or Degree.

Mysore

May, 2009. Register No. 07AUD007

ACKNOWLEDGEMENTS

I would like to devote this work to God. I am, all that I am now because of His never failing love for me.

"I can do everything through Christ who strengthens me". Philippians 4:13.

I would like to thank Dr. Vijayalakshmi Basavaraj for her guidance and tremendous help throughout this study, without her patience and insightful comments it would not have been a success.

I am highly indebted to Sujith sir for his generous help at all times, he has been very patient with my questions and spent tremendous amount of time for this study.

To Dr. Asha Yathiraj and all the faculties of Audiology Department, you have been instrumental in allowing this dissertation to be completed. My heartfelt thanks to you.

My earnest gratitude to Dr. Rindiki, I am overwhelmed with gratitude for all the valuable help and information.

I would like to express my sincere gratitude to Vasanthalakshmi ma'am for helping me out with all the statistical works.

I am deeply indebted to my family- my Mom, Dad, Kimi, Feli, Tatei & Joe for all their prayers, encouragement and support at every step. I am lucky to have a family who loves me as much as you do.

My participants, especially Ruth, without whom this study would have never materialized, my sincere gratitude for your enthusiastic and keen participation.

My wholehearted thanks to Freddy, I am grateful to him for being there with me through the good and bad times, never failing to help and motivate me.

To my friend Tamanna, I couldnt have found a better friend and a wonderful person than you are. My beloved friends, Ramya, Meenakshi, Ridhima, Pallavi, Shuchi, Ismail, Sharat and all my classmates, you have been in my life ever since we started in AIISH. I feel fortunate having you as my friends and I cherish our many years together and hope that we will be able to stay in touch despite the large distances between us.

TABLE OF CONTENTS

CHAPTERS		Page No.
	List of Tables	i
	List of Graphs	ii
1	Introduction	1- 8
2	Review of literature	9- 43
3	Method	44- 52
4	Results and Discussion	53- 68
5	Summary and Conclusion	69- 71
	References	72- 85
	Appendix I	
	Appendix II	
	Appendix III	
	Appendix IV	

LIST OF TABLES

Table No.	Title	Page No.
1	Mean and Standard Deviation of SRT for spondee word lists I & II for right ear and left ear across gender.	54
2	Mean and standard deviations of scores on spondee word lists I & II at different intensity levels across gender.	57
3	Mean and Standard Deviation of SIS for PB word lists I & II for right ear and left ear across gender and across ear.	61
4	Mean and standard deviations of scores on PB word lists I & II at different intensity levels across gender.	64
5	Percent reliability of Spondee list I & II and PB list 1 & 2.	68

LIST OF GRAPHS

Figure No.	Title	Page No.
1	Mean SRT scores for right and left ear for spondee lists I & II.	55
2	Performance Intensity function of spondee lists I & II across the six intensity levels.	58
3	Mean SIS for right ear and left ear for PB lists I & II.	62
4	Mean SIS for lists I & II across six intensity levels.	65

Chapter 1
INTRODUCTION

1.1 Hearing Assessment:

A comprehensive evaluation of an individual's hearing acuity requires several different types of diagnostic techniques. A commonly utilized procedure is pure-tone audiometry.

a) Pure Tone Audiometry: This procedure assesses the thresholds at which a listener is able to detect sinusoidal frequencies. Pure-tone testing is a relatively quick and reliable method to obtain an assessment of an individual's ability to detect specific frequencies. However, to accurately evaluate a listener's ability to comprehend the more complex acoustic signals such as speech, additional auditory tests need to be performed.

b) Speech Audiometry: Speech audiometry is a procedure that is used to evaluate a listener's ability to hear, recognize and understand speech communication (ASHA, 1988; Young, Dudley, & Gunter, 1982). This type of assessment is valuable in the diagnosis of peripheral and central auditory disorders, evaluation of hearing aid candidacy, assessment of hearing aid performance, as well as locating possible lesions within the auditory system. In addition, speech audiometry can be used to validate previously obtained pure-tone average (PTA) results.

The usual measures of Speech Audiometry consist of:

Speech Detection Threshold (SDT): The speech detection threshold is the minimum hearing level for speech at which an individual can just discern the presence of a speech material 50% of the time. The listener does not have to identity the material as speech, but must indicate awareness of the presence of sound (ASHA, 1979).

Speech Recognition Threshold (SRT): The SRT is the lowest hearing level at which a person correctly recognizes the speech stimuli 50% of the time. The terms “speech reception threshold” and “spondee threshold” have been used synonymously with “speech recognition threshold” (ASHA, 1979).

Speech Identification Scores (SIS): The SIS is an assessment of a patient’s ability to identify and repeat single syllable words presented at suprathreshold level (Stach, 1998). They are expressed in percentage of words correct in each list. The other terms used are Word Identification Scores or Speech Discrimination Scores, PB word testing etc. (Katz, 2001).

Most Comfortable Loudness level (MCL): The MCL for speech is the hearing level at which the patient experiences speech material to be most comfortable, that is, where he/she prefers to listen to the speech material. The MCL testing involves adjusting the hearing level of speech until the patient indicates that it is comfortably loud (Gelfand, 2001).

Uncomfortable Loudness Level (UCL): The UCL for speech is the hearing level at which the patient considers speech material to be uncomfortably loud (Gelfand, 2001).

1.2 Diagnostic significance of Speech Audiometry:

Speech Audiometry serves many clinical purposes. The basic purpose is to quantify the listener's hearing level for speech. Speech audiometry is often utilized as a diagnostic tool in determining whether a hearing impairment exists (Bell & Wilson, 2001) and whether the impairment is conductive, sensory, or neural (Egan, 1979; Hagerman, 1984). It is also helpful in diagnosing central and peripheral auditory disorders (Jerger et al, 1983) as well as being used to assess performance and function of cochlear implants (Dowell et al, 1986; Cowan et al, 1997; Sarant et al, 2001). It provides more information than pure-tone audiometry concerning a person's hearing impairment because it analyzes not only residual hearing threshold but also sound distortion, loudness, localization, and speech comprehension (Martin, 2001). It serves as a validity check for the pure tone audiometry. Speech Identification (SI) on the other hand makes it possible to evaluate the functional integrity of the auditory system. The poorer the SI scores, greater is the involvement of sensori- neural mechanism. SI scores can be used to differentiate cochlear pathology to retrocochlear pathology in addition with other test results (Goetzinger, 1972).

The hearing impairment inferred from a pure tone audiogram cannot depict, beyond the gross generalizations, the degree of disability in speech communication caused by a hearing loss. It is logical that tests of hearing function should be performed with speech stimuli. Using speech audiometry, audiologists set out to answer questions regarding patients' degree of hearing loss for speech, the levels required for the speech to be most comfortable, uncomfortable loudness levels, the range of comfortable loudness and perhaps most importantly, their ability to recognize the sounds of speech. Speech-language pathologists and audiologists use findings of speech audiometry in both therapy planning and counseling (Martin & Clark, 2003).

Thus, speech audiometry tests are essential components of any comprehensive audiological evaluation.

The earliest systematic investigation in the area of speech audiometry was in 1920's when Campbell proposed a method of calculating the efficiency of telephone sound-transmitting equipment using nonsense syllables and these nonsense syllables were spoken over communication systems. This way the consonant sounds could be evaluated as to how well or poorly they were reproduced by the system under study. Crandall and his associates refined the syllable choice by randomly obtaining consonant-vowel combinations. These lists were known as 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 "per cent score of articulation" concept, and it is still used today.

The next major development in speech testing came from Harvard Psycho-Acoustic Laboratory (PAL). Egan et al. (1948) developed lists of monosyllabic words for speech testing. The criteria for word and list selection included: (a) lists must be of equal average difficulty; (b) each list must have a composition representative of English speech, i.e., must be phonetically balanced; (c) the words must be in common usage. The resultant test lists are known as the PAL PB-50 Lists.

The PB-50 list was amended as the CID Auditory Test W-22 in 1952. This produced the familiar W-22 recordings with a speaker, who followed the procedure of saying a carrier phrase at a set level and allowing the key word to fall at some presumably natural level relative to the carrier phrase. The W-22 word lists and recordings were developed to overcome some of the shortcomings of the PAL PB-50 lists. The resultant recordings contain easier discrimination tasks than their predecessors.

Harvard University developed spondee lists for measuring Speech Reception Threshold in 1947 and are known as Auditory Test No. 9 and No. 14, each consisting of 42 disyllabic words with a spondaic stress pattern, which are scrambled into 12 orders or lists. The modification of the spondee lists were done at the CID laboratories, as were the original PB lists. From the original 84 spondees, only 36 were kept in the CID W-1 and W-2 lists and recordings. These 36 spondees were scrambled into six lists. In the W-1 recordings, those words that had been found to be too easy were recorded 2 dB lower than the average, and those that had been found to be too difficult were recorded 2 dB higher. The words on the W-1 recording were not homogenous as has been reported by Bowling and Elpern (1961).

Various kinds of speech stimuli have been used to determine the SRT. They are sentences, connected discourse, spondaic words, spoken digits etc. the kind of stimuli used for speech discrimination testing are monosyllables, nonsense syllables, synthetic speech etc. For the estimation of SRT, spondaic words are the most widely used test stimuli and mono words in case of SI tests (Carhart, 1971).

1.3 Need for language specific Speech Tests:

In order for speech audiometry to be a valid and accurate evaluation, individuals should be tested in their native language (Ramkissoon, 2001). Stimuli for Speech audiometry have been developed in various languages. However, such materials may be inappropriate if it is not of the specific regional dialect of the speaker and/or the listener. It has been reported that speech audiometry recordings of a speaker with a non-regional dialect, even if mutually intelligible in ideal listening conditions, may be relatively more difficult for listeners when presented at low-intensity levels or in the presence of noise (Weisleder & Hodgson, 1989). Thus, hearing evaluations using recordings in a non-native dialect may be less valid, especially at presentation levels below 50 dB SPL (Wilson & Moodley, 2000).

Researchers and audiologists have recognized the need for linguistically appropriate diagnostic tools and have developed speech audiometry tests in various languages.

1.4 Need for the study:

High-quality, standardized speech audiometry materials have been developed and used extensively in English. However, for many of the world's languages such materials are more limited or non-existent. Since India is a multilingual country, there is a need to develop the language specific test material. Several Speech Recognition Tests have been developed in Indian languages, such as, Spondaic word lists in Tamil, Telugu and Malayalam by Kapur (1971), SRT Test for adults and children in Kannada by Rajashekhar (1976), Speech Test material in Manipuri by Tanuza (1984), SRT Test in

Bengali by Ghosh (1986), SRT Test in Gujarati by Mallikarjuna (1990) and SRT Test in Oriya by Behera (2004).

Among the Speech Identification tests developed for the adults were Phonetically Balanced (PB) words list in Hindi by Abrol (1971, cited in Nagaraja, 1990), Speech Perception Test in Tamil, Telugu and Malayalam by Kapur (1971), Hindi PB list for Speech Audiometry and Discrimination Test by De (1973), Synthetic Sentence Identification Test in Kannada by Nagaraja (1973), Common Speech Discrimination Test for Indians by Mayadevi (1974), PB Test in Tamil by Samuel (1976), Speech Identification in Manipuri by Tanuza (1984), Speech Identification Test in Bengali by Ghosh (1986) , Speech Identification Test in Gujarati by Mallikarjuna (1990) and Speech Identification Test in Oriya by Behera (2004).

Mizo is a language spoken in the North- Eastern state of Mizoram, with a population of 9 lakhs. This kuki- chin branch of Tibeto- Burman language is unique in the sense that it is tonal in nature, having 5 vowels and 28 consonants, 10 dipthongs, 4 triphthongs, 3 voiced nasals and 3 voiceless nasals. It does not have a script of its own; rather, it shares the Roman script used for the English language.

No speech test material for evaluating the speech recognition threshold and speech identification ability is available in Mizo language. Hence, there is a need to develop and standardize speech material in Mizo for assessing the hearing abilities of subjects who know only Mizo language.

1.5 Aims & Objectives of the study:

The main objectives of the study were:

1) Construction of a bisyllabic word list to assess speech reception threshold in Mizo language
2) Construction of a monosyllabic word list to assess speech identification scores in Mizo language.
3) Standardization/ normalization of the lists prepared.

Chapter 2

REVIEW OF LITERATURE

2.1 Role of Speech Audiometry in hearing assessment:

Speech audiometry constitutes part of examination of auditory system function when a comprehensive evaluation of hearing ability has to be achieved. The purpose of speech audiometry is to evaluate listener's ability to hear, recognize and understand the verbal communication that is encountered in day to day life (Young et al., 1982. ASHA, 1988). Speech audiometry is often utilized as a diagnostic tool in determining whether hearing impairment exists or not (Bell & Willson, 2001), and if exists, whether the impairement is conductive, sensory or neural in nature (Egan, 1979; Hagerman, 1984). It is also helpful in diagnosing central and peripheral auditory disorders (Jerger & Abrams, 1983), as well as used to assess performance and function of cochlear implants (Dowel et al., 1986; Cowan et al, 1997; Sarant et al, 2001). It provides more information than pure tone audiometry as it analyses not only the residual hearing thresholds but also sound distortion, loudness, localization and speech comprehension (Martin, 2001).

In cases of children and difficult-to-test populations, conditioning for speech audiometry is often easier than conditioning for pure tone audiometry. Furthermore, conditioning for speech stimulus often provides carry- over to conditioning for pure tones. In clinical situations, young children usually respond more easily to the presentation of speech materials than to pure tones. As a result, estimates of thresholds for speech recognition are often sought first in children to provide the audiologists guidance in establishing pure tone thresholds (Stach, 1998).

In adults, suprathreshold speech understanding maybe a sensitive indicator of retrocochlear disorder, even in the presence of normal hearing sensitivity. A thorough assessment of speech understanding in such patients may assist in the diagnosis of neurological disease. In elderly individuals, speech audiometry is a vital component in our understanding of the patient's communication function. The degree of hearing impairment described by pure-tone thresholds often underestimates the amount of communication disorder that a patient has, and suprathreshold speech audiometry can provide a better metric for understanding the degree of hearing impairment resulting from the disorder (Stach, 1998).

2.2 Applications of Speech Audiometry:

The ultimate effectiveness of speech audiometric tests will depend to a large degree upon standardization of test materials, recordings, pictures, type and degree of degradation and administering and scoring procedures (House, 1965). Speech audiometric measures are used routinely in an audiologic evaluation and contribute in a number of important ways (Fulton & Llyod, 1975; Stach, 1998) including:

i) Measurement of threshold for speech.

ii) Cross- check of pure tone sensitivity (Chaiklin & Ventry, 1964).

iii) Quantification of suprathreshold speech recognition ability (ASHA, 1979; Wilson & Margolis, 1983).

iv) Assistance in differential diagnosis.

v) Assessment of central auditory processing ability.

vi) Predicting site of lesion.

vii) Predicting performance with amplification (ASHA, 1979, 1988; Wilson & Margolis, 1983).

viii) Estimation of communicative function. (Katz, 2001).

i) Speech Thresholds

A speech threshold is the lowest level at which speech can be detected or recognized. The threshold of detection is referred to as the Speech Awareness Threshold (SAT) or, less commonly the Speech Detection Threshold (SDT). The threshold of recognition is often referred to as the Spondee Threshold (ST) or a speech recognition threshold (SRT). The ST is a measure of the threshold of sensitivity for hearing and identifying speech signals. Even in isolation, the ST provides significant information. It estimates hearing sensitivity in the frequency region of the audiogram where the major components of speech fall, thereby providing a useful estimate of degree of hearing loss for speech. Thresholds for speech measures are highly reliable (Tillman & Jerger, 1959).

ii) Pure tone cross- check

One of the obvious values of the SRT measure is its close relationship with the pure tone thresholds, especially the average of the two best thresholds for the frequencies of 500, 1000 and 2000 Hz. (Carhart, 1946; Fletcher, 1950) That is, if both the pure tone intensity levels and the speech intensity levels are expressed on the dB HL scale, the degree of hearing loss for speech should agree with the degree of hearing loss for pure tones in the 500 through 2000 Hz region. Normally, the speech recognition threshold agrees with the average pure tone thresholds of 0.5, 1 and 2 KHz within ±10 dB (Hagerman, 1979). In practice, speech signals seem to be easier to process and sometimes result in lower initial estimates of threshold than testing with pure tones. In such a case, the pure tone thresholds may actually be suprathreshold and the patient will need to be reinstructed. In pseudohypacusis cases, the speech threshold may be substantially better than the pure tone average. When the audiogram slopes precipitously, it is often useful to compare the SRT to one frequency that has the best threshold, which is often 500 Hz, and can even sometimes be 250 Hz (Gelfand & Silman, 1985, 1993; Silman & Silverman, 1991).

iii) Quantificcation of suprathreshold speech recognition ability

Speech recognition testing is designed to provide an estimate of suprathreshold ability to recognize speech (Silman & Silverman, 1991). In its most fundamental form, speech recognition testing involves the presentation of single-syllable words at a fixed intensity level above threshold. This is often referred to as speech discrimination testing or word recognition testing (Boothroyd, 1968; Stach, 1998). Results of word recognition testing are generally predictable from the degree and configuration of the pure tone audiogram. It is in this predictability that the value of the test lies. If word recognition scores equal or exceed those that might be expected from the audiogram, then suprathreshold speech recognition ability is thought to be normal for the degree of hearing loss. If word recognition scores are poorer than would be expected, then suprathreshold ability is abnormal for the degree of hearing loss. Abnormal speech recognition is often the result of cochlear distortion or retrocochlear disorder. Thus, word recognition testing can be useful in providing estimates of communication function and in assisting in the diagnosis of neurologic disorder (Stach, 1998).

iv) Differential diagnosis

Speech audiometric measures can be useful in differentiating whether a hearing disorder is due to changes in the outer or middle ear, cochlea, or auditory peripheral or central nervous systems. Speech recognition scores are generally expected to be 90% - 100% in normal hearing individuals. The range of speech recognition scores is typically between 80% and 100% with most conductive losses, but has been found to be as low as 60% in glomus tumor, and the range for sensorineural losses is anywhere from 0% to 100%, depending on the degree of loss and etiology (Bess, 1983). In general, speech recognition scores that are ‘abnormally low’ are associated with retrocochlear lesions; however, there is no clear cutoff value for this decision (Johnson, 1977; Bess, 1983). Generally word and sentence recognition scores are lest sffected in the presence of conductive loss and most affected in the presence of neural loss. Both the sensory (cochlear) loss and neural (retrocochlear) loss group of patients produce very wide score ranges, hence, extreme variability in scores (Pendrod, 1994).

v) Central Auditory Processing

Speech audiometric measures also permit us to evaluate the ability of the central auditory nervous system to process acoustic signals. The auditory system, in its vastness of pathways, includes a certain level of redundancy or excess capacity of processing ability. Such redundancy serves many useful purposes, but it also makes the function of the central auditory nervous system impervious in our efforts to examine it. For example, a patient can have a rather substantial lesion of the auditory brain stem or auditory cortex and still have normal hearing and normal word recognition ability. With the use of advanced speech audiometric measures, we are able to measure to identify the presence of neurologic disorder. They are also helpful in that they provide insight into a patient's auditory abilities beyond the level of cochlear processing (Stach, 1998). Monaural and binaural speech recognition ability has the potential to provide insight about the nature and presence of CAPD (Rosser, Valente & Dunn, 2000).

vi) Speech Recognition and site of lesion

Speech audiometric measures can be useful in predicting where the site of lesion might be for a given hearing loss. If a hearing loss is conductive due to middle ear disorder, the effect on speech recognition will be negligible, except to elevate the speech threshold by the degree of hearing loss in the ear with the disorder. Suprathreshold speech recognition will not be affected. If a hearing loss is sensorineural due to cochlear disorder, the speech threshold will be elevated in that ear to a degree predictable by the pure tone average. One exception is in the case of endolymphatic hydrops or Meniere's disease, in which the cochlear disorder causes such distortion that word recognition scores are poorer than predicted from degree of hearing loss. If a hearing loss is sensorineural due to VIIIth nerve lesion, the speech threshold will be elevated in that ear to a degree predictable by the pure tone average. Suprathreshold word recognition ability is likely to be substantially affected. Maximum scores are likely to be poorer than predicted from the degree of hearing loss, and rollover of the performance-intensity function is likely to occur. Abnormal results will occur in the same, or ipsilateral, ear in which the lesion occurs. If the hearing disorder occurs as a result of a brain stem lesion, the speech threshold will be predictable from the pure tone average. Suprathreshold word recognition ability is likely to be affected substantially (Stach, 1998).

Suprathreshold speech recognition scores obtained with monosyllabic PB words have been used as part of a battery of test to differentiate between cochlear and retrocochlear pathology. Johnson, (1977) found that, 30% of 418 eighth- nerve tumor patients had speech- recognition scores less than or equal to 30%. Many (44%) of these eight nerve tumor cases had scores exceeding 60%. On a study done by Olsen, Noffsinger and Kurdziel (1975), none of the patients with eight- nerve tumors had scores below 30%.

vii) Predicting performance with amplification

Speech audiometry plays an important role concerning various procedures for hearing aid evaluation and selection (Carhart 1950; Resnick and Becker, 1963). One reason is because aided and unaided responses to speech can be obtained in a sound field setting. This enables the clinician to evaluate the listener's performance with and without an aid and to compare performance among aids. The degree to which the results are definitive depends on several factors, including: a) the nature of the hearing impairment (audiometric configuration, hearing level for each ear, type of impairment), b) the speech message employed, c) the listening conditions, and d) the goal of the testing session (selection, evaluation, consultation etc.) (Fulton & Llyod, 1975).

Speech audiometry can rule out inappropriate amplification units and delineate hearing aids with adequate amplification to the user. These procedures are far superior to selecting hearing aids on the basis of an audiogram and factory specifications describing the hearing aids' physical amplification characteristics (Fulton & Llyod, 1975).

viii) Estimating communicative function

Speech thresholds tell us about a patient's hearing sensitivity and, thus, what intensity level speech will need to reach to be made audible. Word- recognition scores tell us how well speech can be recognized once it is made audible. Advanced speech audiometric measure tells us how well the central auditory nervous system processes auditory information at suprathreshold levels. Taken together, these speech audiometric measures provide us with a profile of a patient's communication function (Stach, 1998).

2.3 Speech Recognition Threshold (SRT) Tests:

Speech Recognition Threshold (SRT) as defined by ASHA in 1979, is the lowest hearing level at which a person correctly recognizes the speech stimuli 50% of the time.

The terms “speech reception threshold” and “spondee threshold” have been used synonymously with “speech recognition threshold”.

There are several SRT tests that have been developed over the years. The material and the method of construction vary from tests to tests. Various kinds of speech stimuli have been used to determine the SRT. They are sentences, connected discourse, spondaic words, spoken digits etc. In case of SRT tests, spondaic words are the most widely used test stimuli (Carhart, 1971).

1. Harvard or Psycho- Acoustic Laboratory (PAL) Auditory Test No. 9

This test was developed by Hudgins et al. (1947), (cited in O’Neill & Oyer, 1966). The purpose of this test was to measure the threshold of intelligibility of speech. It was an open set test using spondaic words as stimulus and consisted of two lists, each with forty two disyllabic words with equal stress placed upon both syllables in each of the words. A carrier phrase was used and a 1000 Hz calibrating tone was recorded on the phonographic discs where the stimuli were recorded. In the recordings of the lists, the words were attenuated through a range of 24 dB with an attenuation of 4 dB being provided for each group of six words and an interstimulus interval of 6 seconds.

2. Psycho- Acoustic Laboratories Auditory Test No. 14

This test was also developed by Hudgins et al (1947, cited in O’Neill & Oyer, 1966). It is similar to the Auditory Test No. 9. Except in the manner of the recording of the test materials. In this test, all of the spondee words were recorded at one level. No attenuation of the level of the words occurred.

3. CID Auditory Test W-1

This is a modification of the Psycho- Acoustic Laboratories (PAL) lists by Hirsh et al in 1952. They developed the Central Institute for the Deaf (CID) Auditory Test W-1, which is an open set to measure the threshold of intelligibility for speech. From a group of eighty four spondee words in PAL Test No. 9, only thirty six were retained by eliminating those words judged to be either too easy or too difficult. The thirty six spondees were scrambled into six lists recorded at a constant level with a carrier phrase preceding the test word. A 1000 Hz calibration tone was recorded at the same level as the carrier phrase.

4. CID Auditory Test W-2

This test was also developed by Hirsh et al. (1952) and was designed for rapid estimation of the intelligibility threshold. The recorded lists were the same as those used in Auditory Test W-1. However the intensity of the words had been attenuated in such a way that, there was a drop in intensity of 3 dB for every three words. Instead of presenting a whole list or a portion of a list at a fixed intensity or several intensities, this test sweeped through an intensity range of 33 dB by attenuating the level of the test words 3 dB every three words. The carrier phrase was presented at the same level as the first group of three words and remained constant until after the ninth word. Thereafter, the carrier phrase was attenuated at the rate of 3 dB with each new group of three words.

5. *Speech reception threshold testing using sentence stimuli*

This test was developed by Plomp & Mimpen, 1979 (cited in Plomp, 1986). It is an open set test consisting of ten lists of thirteen simple meaningful sentences which was recorded by a trained female speaker. In the selection procedure only sentences with approximately equal chances of correct recognition in noise were retained and were divided into ten lists with equal number of the various phonemes.

6. *Speech Intelligibility in Noise*

The main purpose of this test was to develop a speech material suitable for clinical measurement of speech intelligibility in noise. Hagerman (1982) developed this sentence test in Swedish language. Ten sentences of five words each were chosen to constitute the original list, which was phonetically balanced. A noise was synthesized from the speech material to produce exactly the same spectrum of speech and noise. The pause between the sentences was 7 seconds and the total time of each list was 110 seconds. This material was intended to be used especially for hearing aid evaluation, but also to measure the patient's speech discrimination in noise with earphones during hearing aid fitting.

7. *Hearing in Noise Test (HINT)*

The HINT test was developed by Nilsson, Soli and Sullivan (1994) to provide a reliable and efficient measure of speech reception thresholds for sentences in quiet and in noise.

It consist of twenty five phonemically balanced lists of ten sentences that had been normed for naturalness, difficulty and reliability, and consist of noise that is filtered to match the long- term average spectrum of the sentences. HINT sentence lists are presented at 65 dB (A) in quiet, or at a fixed SNR (+10 dB, +8dB, +5dB or +0dB signal-to- noise ratio).

8. *Sentence SRT Test in Dutch*

This test was developed for measurement of speech reception threshold by Versfeld, Daalder, Festen and Houtgast (2000). The material used was sentences which were recorded by two male and two female speakers. It has two sets each comprising of thirty nine lists of thirteen sentences. In total, 1014 new sentences were generated that were equally intelligible when presented in speech shaped noise at equal rms level.

Though these tests differ in terms of the speech stimuli used and the method to establish SRT, the overall applicability remains the same.

2.4 Speech Recognition Threshold Tests developed in India

Most of the SRT tests in Indian languages are the outcome of post graduation dissertations submitted as part fulfillment for their degrees.

1. *SRT Test material in English for Indians*

Swarnalatha (1972) was the first one to standardize speech material in English for use on the Indian population. Eighty four words from the PAL Auditory Test No. 9 and No. 14 were used for the adult population and fifty seven words from the children's spondee list were used for children. The final spondee list for adults and children had fifty and twenty five words respectively. All the test items were recorded preceded by a carrier phrase 'say the word'. A time interval of 5 sec was allowed for the subject to respond. The subjects were instructed to respond orally.

2. *SRT test for adults and children in Kannada*

This speech test was developed by Rajashekhar (1976). The test materials were obtained from the Central Institute of Indian Languages (CIIL) and consisted of 104 bisyllabic or polysyllabic words which could be picturable. The words were subjected to familiarity and intelligibility check. Finally forty words were randomly selected and put into two lists. Pictures were used to test the familiarity with children. The words were recorded with an interstimulus interval of 5 seconds. Standardization of the speech tests was done with adults and children.

3. *SRT test in Manipuri*

Tanuza (1984) aimed at construction and standardization of SRT test material in Manipuri language. After familiarity test, eighty polysyllabic words were chosen for the construction and four lists were developed. The polysyllabic word list contained twenty items each. All the test materials were tape recorded and spoken with an interstimulus interval of 5 seconds. For the standardization of the test material, four lists of twenty five items each were used. The presentation level was 0, 5 10, 15 20 and 25 dBHL with respect to 0 dB SRT. Each list was randomized into six lists and each randomized list was presented at only one intensity level. The lowest level at which the subjects repeated correctly 50% of the test items was considered as the SRT level.

4. *SRT test in Bengali*

The test was given by Ghosh (1986). The objective of the study was to construct and standardize speech materials in Bengali language to facilitate the speech audiometry procedure. Sixty polysyllabic words were chosen for assessing the SRT. All the test items were recorded preceded by a carrier phrase. For the standardization of the test material, the polysyllabic words were divided into three lists consisting of twenty words and each list was randomized into six lists to overcome practice effect. The level of presentation for the list were 0, 5, 10, 15, 20 and 25 dBHL ref. 0 dBHL.. The level at which the subject repeated correctly 50 % of the test items were taken as the SRT level.

5. *SRT test in Gujarati*

Mallikarjuna (1990) developed a list of sixty spondee words in Gujrati language, for which the articulation as a function of intensity was determined. A positive correlation of 0.73 and significant at 0.01 level was established with spondee list in English for Indian population (as standardized by Swarnalatha, 1972).

6. SRT test in Oriya for Adults

A speech reception test in Oriya was developed by Behera.S (2004) for the adult population. The first stage of the procedure involved construction of test material for SRT test and the second stage involved obtaining normative data for the same. For SRT testing, two lists were developed consisting of 25 bisyllabic words each. Familiar bisyllabic words representing all phoneme of the language were selected. Forty adults participated in obtaining normative data. Here, the subjects were tested for establishment of SRT by employing the Martin & Stauffers (1975) method for obtaining SRT.

7. *Speech Perception Test in Tamil and Telugu*

It was developed by Kapur, 1971, for adults. The objective was to construct auditory word lists (spondees) for the construction of SRT test in both the languages and to establish reliability and validity. Except for the nature of the material used in the construction of these tests and methods of selection, methodology for tests in both the languages was similar. Speech Audiometry tests in Tamil and Telugu was done by selecting words which were very common. The recorded 200 disyllabic words were presented at threshold +4, +2, 0, -4, and -6 dB relative to the pure tone average thresholds of the subjects. Good correlation was obtained between the SRT and the pure tone threshold.

However, in Tamil language, the list failed to represent all the sounds which occur in the language.

2.5 Speech Identification (SI) Tests

Word recognition scores usually are obtained in each ear separately with what are called “phonetically balanced” but are actually “phonemically balanced” (PB) monosyllabic words.

In the past many authors have used the term ‘discrimination’ for ‘identification’. However, discrimination refers to differentiation among stimuli as ‘same’ or ‘different’. Hence speech discrimination is considered an inappropriate term. (Olsen & Matkin, 1979).

The following section gives some of the speech identification tests reported in literature, the ones developed in India are given separately.

Fletcher and Steinberg (1929) gave one of the earliest test to measure speech identification. Consonant-vowel-consonant (CVC) monosyllabic nonsense syllables were chosen out of which ten lists of ninety syllables were formed. Twenty two introductory sentences were used to make the test more like connected speech. This open set format test was presented through monitored live voice and twenty two different introductory phrases were used as carrier phrase. Through this study on nonsense syllables, Bell's Laboratory developed the necessary specifications for quantifying the quality of speech sounds in telephones to improve speech intelligibility.

Fry and Kerridge, 1939 (cited in Olsen & Matkin, 1979) described five lists of twenty five CVC words each, called word test for deaf people. Each list contained all but four sounds of English spoken language. The list of words were to be spoken by a friend or relative in a quiet room at an agreed upon distance. The listener closed his/her eyes and responded to each word. Fry and Kerridge advised that any person whose whole score was less than 35% should be tested with a list from their sentence to allow the listener the advantage of context.

Fry and Kerridge, 1939 (cited in Olsen & Matkin, 1979) developed Sentence tests for deaf people prepared sentence tests for patients who obtained scores of 35% or less on their word lists. These consisted of five lists of sentences. There were twenty five sentences per list with four to seven words per sentence.

The PAL Auditory test No. *9* was developed by Hudgins et al. 1947 (cited in Mendel & Danheur, 1997) to determine listener's ability to hear simple sentences in the presence of interfering noise. It uses a multiple choice format. It consists of two spondaic- word lists, each list containing 42 spondees. These lists were used clinically for measuring the patient's speech threshold. The criteria employed to construct the spondaic- word lists developed by Hudgins et al. (1947) were:

a) Familiarity with respect to vocabulary
b) Phonetic dissimilarity, so one spondaic word would not be easily confused with another spondaic word
c) Normal sampling of English speech sounds in everyday conversation is essentially the same as the proportion of occurrence of speech sounds in the spondaic word list
d) Homogeneity with respect to audibility.

PAL PB- 50 word list was then developed consisting of twenty lists of fifty words each included with reference to the following criteria:

Equal average difficulty, equal range of difficulty, equal phonetic composition, representative of English speech and words in common usage.

Various tests for Speech Identification constructed over the years:

1. *CID W-22 Word List*

Hirsh et al. (1952) constructed PB word lists for use with adults. These lists were constructed with the aim of increasing average familiarity of words, in comparison with the Harvard PB lists (1947) and by increasing the homogeneity with respect to audibility. This homogeneity was improved by eliminating the "easy" words (words missed once or less by all listeners when the spondee lists were presented at +4, +2, 0, -2, -4, and -6 dB SL) and "hard words" (words missed five or more times by all listeners when the spondee lists were presented at +4, +2, 0, -2, -4, and -6 dB SL with respect to the threshold for the PAL Test 9). The words chosen were monosyllabic, and grouped into four lists of fifty words each, all of which were phonetically balanced. The test was standardized on a small group of fifteen subjects. Hirsh et al. (1952) had not reported that mean scores obtained at the different levels for the four lists. The scores obtained could be derived only from the articulation curve reported by them. To demonstrate the significance of difference among lists and among levels, no statistical procedure was employed. The conclusion was based only on the articulation function.

2. *CID everyday sentences*

Davis & Silverman (1978) published sentences that were developed at Central Institute for Deaf (CID) to represent 'everyday American speech'. It is made up of 10 sets of 10 sentences with 50 key words in each set; they have been named the CID everyday speech sentences. The scoring for this open- response test is based upon the number of key words correctly identified. There is no known recorded version of this test.

3. *Multiple Choice Intelligibility Test*

This test was developed by Black (1957). This closed- set multiple choice format test consisted of twenty four multiple choice lists, each consisting of twenty four test items. They were constructed by:

- Selecting materials from a master population of words
- Collecting the error responses for the words through written down tests
- Assembling of trial form of multiple choice list and answer form and
- Assembling of test forms C & D and A& B.

The purpose of developing this test was to determine a speaker's ability to be heard correctly, to determine a speaker's intelligibility under quiet and noise conditions, to measure listener's efficiency as a reflection of listener's performance in communication situations.

4. *Consonant- Nucleus- Consonant (CNC) word lists*

Lehiste & Peterson (1959) developed a new monosyllabic word test. They introduced the concept of phonemic balancing. They defined the middle vowel as the "nucleus". In the construction of the test, 1263 monosyllabic words of the CNC type were selected. The frequency of occurrences of each initial consonant, vowel nucleus and final consonant was determined and incorporated into the construction of ten lists of fifty words each.

5. *PAL S-1 Sentence Test*

The PAL S-1 sentence test developed by Davis & Silverman (1960) was an open response test designed as a suprathreshold measure of speech intelligibility (recognition). They were developed based on the rational that sentences can be used when patients have extreme difficulty with monosyllabic word tests, or when the audiologist wishes to have a better approximation of how well a patient understands contextual material. There were two equivalent lists of 20 sentences and in each sentence there were five key words (four monosyllables and one bisyllable). The listener writes down or repeats each sentence. Scoring is based on recognition of the five key words in each sentence.

6. *NU Auditory Test No. 6*

In an attempt to improve phonemic equivalency and equal word familiarity from list to list, Tillman and Carhart (1966) refined the CNC lists and produced four equivalent lists. They were recorded on a tape by a single male talker and were designated lists 1, 2, 3 and 4. Each list was rerecorded into several scramblings with a letter designation (a, b, c etc.). The recordings were called the Northwestern University Auditory Test number 6.

7. *Speech Perception in Noise (SPIN)*

Kalikow (1977) developed the recorded SPIN test to assess for speech understanding in noise. This open- set response test contains eight sets of fifty sentences on one track of an audio tape. Only the last word in each sentence is the test item, resulting in 200 test words. Each sentence has five to eight words. The other track contains the babble of 12 voices. The test sentences and the babble can be mixed at various speech-to- babble ratios. Half of the sentences contain high- predictability items and half contain low- predictability items, based upon contextual, syntactic, and prosodic cues.

8. *Revised Speech Perception in Noise Test*

Bilger (1984) found that only six of the eight forms were equally difficult. He rerecorded the six equally difficult forms. This revised SPIN test is often used to assist in hearing aid selection (Katz, 2002).

9. *Maryland CNC*

Causey et al. (1984) described the Maryland CNC test, named after the University of Maryland where it was developed. These investigators reemphasized that phonemes preceding and following the middle nucleus sound affect the production, and hence, the acoustics, of that nucleus sound. Furthermore, they pointed out that the acoustic parameters of consonant sounds are influenced by the transitional shifts which occur where a vowel and a consonant join. This phenomenon is referres to as “coarticulation” and the authors believe this is a factor that should be controlled when obtaining word recognition scores.

They took 500 of the original NU CNC words and imbedded each of them in the phrase “Say the (test word) again” because they believed the same “acoustical surround” should be used for the best control of the co- articulation factor. Ten PB lists, each with 50 different words, were recorded on tape by a single trained male talker. Performance intensity functions were completed for a group of 60 college- aged subjects and a group of 40 hearing impaired patients ranging in age from 30- 74 years. Six of the ten recordings were deemed the most equivalent and were rerecorded on CDs. This test is gaining popularity in the United States.

10. *Rhyme Test*

This test was developed by Fairbanks (1958) on the basis of the multiple choice test. It was one of the first commercially available non- open- response word recognition tests. It was designed to emphasize the auditory phonemic factors and to minimize the linguistic factors. Five lists of rhyming monosyllables with fifty items per list were used. Only the initial consonant differed in a set of five rhyming words, but the initial consonants were not given on the response sheet. Hence, the response task of the listener was to write the initial consonant for the stem provided in the answer sheets. This rhyme test was labeled as a phoneme recognition test, but is actually a consonant recognition test.

11. *Modified Rhyme Test*

The Rhyme test was modified to a true closed set response test by House et al. (1965). It contains six 50 monosyllabic word lists, each list requiring the identification of 25 initial and 25 final consonants.

12. *California Consonant Test (CCT)*

Owens & Schubert (1972) developed a four- alternative forced- choice closed- response test called the California Consonant Test. In this test, the listener is given an answer sheet and told to select one four words following each presentation. One of the words is correct and the remainders are foils. The total test list contained 100 CVC items, which are not phonemically balanced were arranged into two scrambling to produce two test list. Instead, the selected items are based upon phoneme recognition errors made by subjects with hearing loss. This test was designed to be sensitive to the discrimination problems of patients with high- frequency hearing losses. These patients show some difficulty.

13. *High Frequency Response Test*

Other word recognition tests have been developed to determine phoneme recognition errors in patients with high frequency sensory neural loss. They include those of Gardner (1971) and Glaser (1974). The lists for these tests are biased with an abundance of high frequency phonemes, and the tests are particularly useful in evaluating improvements in word recognition with various hearing aids/ assistive listening devices in place.

14. *PBK Test*

The Phonetically Balanced Test of Speech Discrimination for Children (PBK-50) uses an open response type task. Based upon what was considered familiar monosyllabic words for children entering the first grade, three equivalent, 50- item word lists were developed (Haskins, 1949). It is recommended that the PBK word lists not be used for children under 6 years of age unless there is assurance that their receptive listening skills are as good as typical normal hearing 6 year olds.

15. *WIPI Test*

The Word intelligibility by Picture Identification (WIPI) test by Lerman et al. (1965), revised by Ross and Lerman (1970), uses a closed- response task. It consists of 25 pages and/ or plates with six pictures on each. For each set, there are four “point to” commands with an appropriate picture for each command. Each set also contains two foils to decrease the possibility of elevated scores due to guessing. The child is presented with a series of cards, each of which contains six pictures. Four of the six pictures are possibilities as the stimulus item on a given test, and the other two pictures on each card (which are never tested) act as foils to decrease the probability of a correct guess. The children have to indicate which picture corresponds to the word they believe they have heard. This procedure is very useful in working with children whose discrimination for speech cannot otherwise be evaluated, provided that the stimulus words are within the children’s vocabularies. Ross and Lerman (1970) suggest that the WIPI test is appropriate for children as young as 4 years of age.

16. *The Northwestern University Children’s Perception of Speech (NUCHIPS) test:*

This test was developed by Elliot & Katz in 1980. This test is similar to WIPI except that, in this the child is presented with a series of four picture sets, including 65 items with 50 words scored on the test.

17. *Synthetic Sentence Identification test*

This test developed by Jerger et al. (1968) involves a set of ten synthetic sentences. Each sentence contains 7 words, with a noun, predicate, object, and so on, but carries no meaning. The sentences were recorded on CD or audio tape, and patients show their responses by indicating the number that corresponds to the sentence they have heard. Some sentences are more difficult than the others. Because early experimentation showed that synthetic sentence identification is not sufficiently difficult when presented in quiet, a competing message of connected speech is presented, along with synthetic sentences, and the intensity of the competing message is varied.

2.6 Speech Identification Tests developed in India

1. *PB word list in Hindi*

This was constructed by Abrol, 1971 (cited in Nagaraja, 1990). He analyzed 800 commonly used words of Hindi for syllabic constructions. The majority of the words were found to have a CVC structure. Two lists of 50 words each were prepared based upon the frequency counts and familiarity of the words. Thirty normal subjects with SRT ranging from 10-30 dB were studied by presenting the material with a carrier phrase "say the word" at 10 dB above the presentation level of the test word. At 10 dB above the SRT, slightly more than half of the population repeated 90% of words and at 30 dB above SRT, all the subjects repeated all the words presented to them. The optimum for Hindi PB was tentatively kept at 20 dB above SRT.

2. *Speech Perception Test in Tamil, Telugu and Malayalam*

Developed by Kapur, 1971 (cited in Nagaraja, 1973), except for the nature of the material used in the construction of these tests and methods of selection, methodology for tests in all these three languages were similar.

In Tamil language, the list failed to represent all the sounds which occur in the language. Speech Audiometry tests in Malayalam was done by selecting words which were very common, for both SRT and PB word lists as very few monosyllabic words were available. In Malayalam language, disyllabic words were used for both SRT and PB word lists as very few monosyllabic words were available in the language. The recorded 200 disyllabic words were presented at threshold +4, +2, 0, -4, and -6 dB relative to the pure tone average thresholds of the subjects.

3. *Speech Discrimination test material in English for Indians*

Swarnalatha (1972) developed and standardized speech discrimination test material in English for Indians apart from developing an SRT test. For the speech discrimination test for adults, 200 monosyllabic words by PAL and another 200 monosyllabic words from CID Auditory Test W-22 were combined and used. For developing the children's discrimination test, 150 monosyllabic words from the KPB word lists were used. The procedure of familiarity was done as in the SRT test.

No data on frequency of occurrence of phonemes in Indian English was available. Hence, the relative frequency of occurrence of phonemes was obtained by analyzing telephone conversation (Fletcher, 1965) and this served as a basis for preparing monosyllabic words. Finally 4 monosyllabic word lists, which were phonetically balanced for adults and two lists of equal familiarity for children.

4. Hindi PB list for Speech Audiometry and Discrimination Test

De (1973) developed spondee and PB word list in Hindi. He finalized 6 lists of 50 words each, all the lists were phonetically balanced. Standardization of the test material was obtained by carrying out clinical tests with these tests on normal Hindi speaking adult subjects.

5. *Development of Synthetic Sentence Identification (SSI) test in Kannada*

Nagaraja (1973) constructed SSI using most commonly used words in Kannada language. Ten first order sentences and ten second order sentences were constructed. To make the task more difficult, these sentences were recorded with a continuously competing speech message. To compare the performance of four subjects of different audiogram pattern in terms of area, Performance Intelligibility function curves were drawn for PB and SSI scores obtained at different intensity levels.

6. *Common Speech Discrimination Test for Indians*

Mayadevi (1974) attempted to construct a speech discrimination test that could be used with the speakers of all Indian languages. She chose the common monosyllable of CV combination found common in Indian languages. The lists of monosyllables were phonetically balanced. The test materials were tape recorded and the level of presentation was kept constant i.e., at definite sensation level above the individual's pure tone average. An intelligibility test was carried out by presenting six lists at different levels to 10 normal subjects. The validity and reliability was checked on normals and clinical groups.

7. *PB Test in Tamil*

Samuel (1976) developed meaningful, familiar, monosyllabic word lists, which were phonetically balanced. The concurrent validity of the test was checked by presenting English PB list to normals having knowledge of English language and analyzing the scores. Normals obtained optimum scores at 35 dBSL (re SRT). All the four lists prepared were found to be essentially equivalent and can be used interchangeably.

8. *Speech Identification in Manipuri*

The study was done by Tanuza (1984). To construct the list, monosyllabic words were collected from phonetic books, magazines, books and normal conversational speech. A familiarity test was done and based on this 100 monosyllabic words were chosen for the construction. Four monosyllabic word lists were divided which contained 25 items, five scrambling were made of the list to avoid practice effect. The monosyllabic words were not phonetically balanced as studies were not available.

9. *Speech Identification test in Bengali*

Ghosh (1986) developed this test. He collected 75 monosyllabic words for assessing speech discrimination ability after a familiarity rating. For standardization, the words were divided into three lists and each of the list was further randomized into five lists and presented at (5, 10, 20, 30, 40) dBSL (re SRT).

10. *Speech Identification Test in Gujrati*

Mallikarjuna (1990) formed three lists of fifty phonetically balanced words. The speech discrimination scores were obtained at various sensation levels above the SRT.

11. *High Frequency Speech Identification test for Hindi and Urdu speakers*

This test was developed by Ramachandra (2001) with the objective to develop a list of meaningful, familiar, high frequency monosyllabic CVC words common to Hindi and Urdu speakers for the establishment of speech identification scores. 50 high frequency words which were rated as most familiar were selected. The first category consisted of high frequency phonemes in initial consonant position of the word and the second category consisted of high frequency phoneme in the final consonant position of the word.

The common high frequency monosyllabic word lists for Hindi and Urdu speakers was both valid and reliable.

12. *A High Frequency Kannada Speech Identification Test*

Mascarenhas (2002) developed a speech identification material in Kannada for testing adults with a sloping hearing loss.

All of these tests developed to assess the speech identification ability had some similarities in terms of type of stimulus used, phonetic balancing, familiarity rating, carrier phrase, instruction, presentation mode, presentation level, response mode and scoring.

13. *Speech Identification Test for Adults in Oriya*

Developed by Behera (2004), familiar monosyllabic words were chosen and 2 lists, consisting of 50 monosyllabic words. Phonetic Balance was maintained in each list.

2.7 Factors affecting Speech Audiometry:

i) *Familiarization and practice effect*

The purpose of familiarization is to ensure that the patient knows the test vocabulary and is able to recognize each word auditorily, and that the clinician can accurately interpret the patient's responses (ASHA, 1988). Tillman and Jerger (1959) found that the practice effect on the SRT is slight (1.1 dB, on average) whereas familiarization with the test list prior to the administration of the test improved the SRT by 4-5 dB on average. Previous findings suggest that practice has a negligible effect on the SRT and that familiarization is necessary to prevent inter-subject differences in the SRT resulting from prior knowledge of the test vocabulary.

ii) *Recorded/ live- voice representation*

The use of recorded speech standardizes the composition and presentation of the speech stimuli and controls for signal intensity. The disadvantages of a recorded presentation include lack of flexibility, which may be important for the difficult-to-test population, and deterioration of phonographic and tape recordings resulting in signal distortion and the introduction of noise.

The advantage of live voice is the flexibility in presentation of the stimuli, especially with regard to the inter-stimulus interval. The disadvantages of live voice include lack of control of signal test intensity despite peaking each syllable at 0 on the VU meter and loss of standardization (ASHA 1979, 1988). Several investigators have reported that the SRTs obtained using monitored live voice is reliable (Beattie, Forrester & Ruby, 1976; Carhart, 1946; Creston, Gillespie & Krahn, 1966).

iii) *Carrier phrase use for SRT testing*

There are equivocal findings on the use of carrier phrase for SRT testing. Some studies found statistically significant advantages for speech recognition scores obtained with a carrier phrase (Gladstone & Siegenthaler, 1971; Gelfand, 1975) whereas others have found no significant differences (Martin, Hawkins & Bailey, 1962; McLennan & Knox, 1975).

iv) *Initial testing level*

Routine speech recognition testing is often done at one hearing level for each ear. Some audiologists add a second measurement at high levels to screen for the possibility of rollover, and testing of speech recognition at more than one level is strongly encouraged. Even though many audiologists perform routine speech recognition testing at Most Comfortable Level (MCL), this is not desirable because the MCL is actually a range rather than a level, and the highest speech recognition score is obtained at levels significantly above the MCL. (Clemis & Caver, 1967; Ullrich & Grimm, 1976; Dirks & Morgan, 1983).

For speech discrimination scores, the literature suggests several levels and procedures for establishing these levels. The levels range from 25 to 50 dB above the speech reception threshold (Carhart, 1946; Newby, 1972).

v*) Testing techniques*

It is generally accepted that the SRT is the lowest hearing level at which a patient can repeat 50% of spondee words, but there are many ways to find this point and no single technique is universally accepted. Most SRT testing methods share a number of common features even though their specific characteristics can vary widely. In general, the differences between SRTs obtained with the various test methods tend to be either not statistically significant or too small to be of any major clinical relevance. SRTs are about 2.7 to 3.7 dB better (lower) with the ASHA (1998) method compared with the ASHA (1979) procedure, but the 1979 method has a slight edge with respect to agreement between the SRT and pure tone averages (Huff & Nerbonne, 1982; Jahner, Schlauch, & Doyle, 1994).

vi) *Phonetic/ Phonemic balance*

The concept of phonetic/ phonemic balance played a major role in the development of many speech recognition tests. However, phonetic/ phonemic balance has been found to have little practical impact on the outcome of speech recognition tests, and its clinical relevance is questionable (Aspinall, 1973; Bess, 1983).

vii) *Whole-word versus phonemic scoring*

Speech recognition tests that use words are usually scored on a whole word basis. Whole word scoring reflects the patient's correct reception of the intended word, but also misrepresents how well the patient is able to make use of acoustical cues of speech. An alternative approach is to score word recognition on a phoneme- by- phoneme basis (Gelfand, 1993; Olsen, Van Tasell, & Speaks, 1997).

Compared to whole word scoring, the use of phonemic scoring -

- Provides a more precise and more valid measure of the correct reception of the acoustic cues of speech
- Improves reliability by maximizing the number of scorable items
- Makes it possible to obtain meaningful scores from patients whose whole word scores would have been zero
- Gives a better idea of which speech sounds are misperceived
- Minimizes the effects of non acoustic factors such as word familiarity, word level predictability, context and differences between word lists.

Comparing word scores and phoneme scores makes it possible to estimate the benefit to speech recognition provided by taking advantage of lexical information (Nittrouer & Boothroyd, 1990; Olsen, Van Tasell, & Speaks, 1997).

viii) *Test size*

Even though most standardized speech recognition tests include 50 monosyllabic words, many attempts have been made to reduce the test size to 25 or even fewer words (Bess, 1983). The problem with reducing test size is that reliability depends on the size of the test. Shortening a test also makes it less reliable because the variability of speech recognition scores is largely defined by the binomial distribution (Boothroyd, 1968; Hagerman, 1976; Raffin & Thornton, 1980; Gelfand, 1993, 1998). Specifically the variability of a test score can be described in terms of its standard deviation, which depends on the percent correct and the number of scorable items in the test.

ix) *Foreign language influences and implications*

Speech audiometry involves material that is inherently linguistic in nature, so the results may be influenced by such factors as differences in phonology and morphologic rules between languages, and are exacerbated by word familiarity effects. Hence, non native speakers of the language typically obtain lower scores on English speech recognition tests than do native speakers of the language. (Gat & Keith, 1978). The perfect solution is for every patient to be tested in his native language by an audiologist who is also a native speaker or at least a fluent speaker of that language.

Nominal facilities are available for assessment of hearing impairment in the state of Mizoram, the Government Civil Hospital has facilities for pure tone audiometry and immitance testing which is being carried out by the ENT professionals, due to lack of audiologists in the field. In the year 2008, two private clinics have been established in the capital city by audiologists, the services offered include audiometric evaluation, immittance evaluation and hearing aid trial.

From the review of the literature on speech perception tests developed, it is evident that attempts have been made to develop tests in different languages. Due to unavailability of speech perception test in Mizo language, and the immense need that exists for the same, an attempt is being made to develop a test material for speech audiometry, taking into account, the various variables which could affect speech audiometry.

Chapter 3

METHOD

The study was conducted with an aim to develop and standardize spondees and phonetically balanced word lists for speech recognition and identification tests in Mizo language.

The study was carried out in two stages:

Stage I: Construction of test material for the Speech Recognition Threshold Test and Speech Identification Tests.

Stage II: Standardization of the test material using Mizo speaking adult subjects.

3.1. Stage 1: Construction of the test material-

a) Obtaining familiar, equally stressed bisyllabic words and monosyllabic (cvc) words
b) Constructing lists of bisyllabic and monosyllabic words.

a) Procedure of familiarity:

In the absence of documents on phonemic and morphophonemic counts in Mizo language, familiar bisyllabic and monosyllabic words were selected randomly from different sources like magazines, newspapers, books and telephonic conversations of individuals fluent in the language. From a corpora of about 1, 00,000 words, familiar 713 bisyllabic and 414 monosyllabic words were selected randomly. A linguist who is familiar with Mizo language was consulted to confirm whether the bisyllabic/ monosyllabic words selected were indeed bisyllabic or monosyllabic respectively.

To further ensure familiarity of the words selected, they were given to 20 normal adults in the age range of 18 years to 40 years, whose mother tongue was Mizo. The subjects were asked to rate the words on a three-point scale of familiarity (i.e., most familiar, familiar and unfamiliar). Words rated as ‘most familiar’ to 90% of the subjects were selected for inclusion in the test lists.

b) Construction of the lists:

For the SRT testing, two lists (List I and II) were developed consisting of twenty-five bisyllabic words each. It was ensured that each list has all the phonemes of the language and equal stress maintained on both the syllables of the bisyllabic word. Perceptual judgment of the equality of the stress was done by a linguist and a speech language pathologist. The words which were classified as having equal stress on both the syllables were then considered for the lists.

For the SI testing, two lists (List 1 and List 2) were developed consisting of fifty monosyllabic words each. The phonemic balance in the word lists were done based on the frequency of occurrence of phoneme in Mizo. Due to unavailability of documents on frequency count of occurrence of a phoneme in Mizo language, the frequency of occurrence of each phoneme in the same corpora, which consisted of the book “Nitin Lalpa Kebulah” authored by Lalnghinglova (2003), with approximately 99,400 phonemes, “Vanglaini” newspaper, which has approximately 6,080 phonemes and the corpora from telephonic conversation was approximately 700 phonemes was done. The number of times each phoneme occurred in the corpora was tallied and counted, and then their percentage of occurrence was calculated and ranked in the order of decreasing frequency. The ranking was divided into 4 equal quadrants. The 1st quadrant consisted of sounds occurring very frequently, the 2nd and 3rd quadrants consisted of sounds occurring frequently and the 4th quadrant consisted of sounds not occurring frequently. The relative frequency of occurrence of phonemes in Mizo language was kept in mind while choosing the words with different phones in the list. Thus the phonemic balance was maintained in each of the lists.

The final lists of spondee words and monosyllabic words prepared are given in Appendix I, II, III & IV.

3.2 Recording procedure:

The lists were recorded with inter- stimulus interval of 6 seconds and normalized for amplitude using Adobe Audition version 2.0 in a sound treated room. The recording was made by an adult female speaker whose mother tongue is Mizo, experienced in the monitored live voice technique of speech audiometry. All the test items were recorded preceded by a carrier phrase “Sawi rawh le” (Say the word). At the beginning of each list, a 1000 Hz calibration tone was recorded. The level of the tone was adjusted so as to produce a 0 VU deflection on the meter of the audiometer.

Stage II: Obtaining normative data

One hundred (100) adults in the age range of 18 years to 40 years (mean age 25 years) were selected for obtaining normative data. The subjects who participated in the familiarity rating were excluded from this group. The subjects met the following criteria to be considered for the study:

i) Hearing sensitivity within normal limits i.e. air conduction thresholds less than or equal to 15 dBHL at all frequencies from 250 Hz to 8 KHz for both the ears.

ii) Have normal middle ear functioning.

iii) Do not have any history/presence of otological problems.

iv) Do not have any speech problems

v) The mother tongue and language spoken at home is Mizo, a language spoken in the state of Mizoram, in India.

3.3 Instrumentation:

i) A calibrated two channel diagnostic audiometer (OB 922), with TDH- 39 headphones housed in MX- 41/ AR ear cushion, calibrated in accordance with ANSI, 1996 S3.6 was used for initial hearing assessment as well as to carry out speech audiometry.

ii) A calibrated GSI- Tympstar immitance meter to ensure normal middle ear condition in the subjects.

iii) Philips CD player, which fed the recorded speech material to the tape input of the audiometer which in turn was fed to the earphone (TDH- 39) housed in MX-41/AR cushions.

3.4 Test environments:

The test was carried out in a sound treated double room situation. The ambient noise levels were within permissible limits, as recommended by ANSI, 1991 S3.1 standards.

3.5 Test procedure:

i) All the subjects were subjected to routine audiological testing by obtaining air conduction and bone conduction thresholds for the frequencies 250 Hz-8000 Hz and 250 Hz-4000 Hz respectively using modified Hughson & Westlake procedure (Carhart & Jerger, 1959). Only those who obtained normal hearing were selected for further evaluation.

ii) Tympanometry for 226 Hz probe tone was done for all subjects. Ipsilateral and contralateral acoustic reflex thresholds were obtained for 500Hz, 1 KHz, 2 KHz and 4 KHz for all the subjects.

3.6 Instructions:

The subjects were given the following instructions in Mizo language:

Instruction for SRT testing: "You will hear a word after the sentence, "Sawi rawh le" through your headphones. Listen carefully to each word and repeat them. The words will get softer. If you are not sure of the word, you can guess the word".

Instruction for SI testing: “You will hear some short words through the headphone. Listen carefully to each word and repeat them”.

3.7 Normative data for SRT test material

Using the material developed for SRT, each of the subjects was tested for the following:

a) Establishment of SRT
b) Performance intensity (articulation gain) function of the spondee word lists.

a) Establishing SRT:

The ASHA (1988) method for SRT determination was followed to evaluate the speech recognition threshold. The procedure is as follows:

Preliminary phase to obtain starting level:

i) The hearing level was set to 30-40dB above the estimated SRT and one spondaic word was presented to the client. If the response was correct, then the level was descended in 10 dB decrements, presenting one spondaic word at each level until the client responded incorrectly. If the client did not respond correctly to the first spondaic word at the first level, the level was increased in 20 dB steps until a correct response was obtained. Then the 10dB decrements were initiated.

ii) When one word was missed, a second spondaic word was presented at the same level. This process of descending in 10dB steps was continued until a level was reached at which two consecutive words were missed at the same hearing level.

iii) The hearing level was increased by 10 dB (above the level at which two spondaic words were missed). This defined the starting level.

Test Phase:

i. Five spondaic words were presented at the starting level and at each successive 5dB decrement.

ii. This was continued if five out of the six words were repeated correctly.

If this criterion is not met, the starting level was increased by 4- 10 dB.

iii. The descending series was terminated when the client responded incorrectly to five of the last six words presented.

Then thresholds were calculated for both the ears, as per ASHA (1988) recommendations.

b) Performance intensity function of the spondee word lists:

The word recognition of the spondee word lists were established at different intensities, starting from 0 dBSL to 10 dBSL progressing in 2 dB steps. The subjects were instructed to repeat the test words and the responses noted down. It was ensured that all the subjects heard the different intensities. At each intensity, both lists (I&II) was presented. At each intensity level, the order of words in the lists was randomized at each intensity level in order to avoid familiarity effect. The average percentage correct scores for both the lists (I&II) were plotted as a function of intensity. This is called the Performance Intensity function.

3.8 Normative data for Speech Identification (SIS) material:

Using the material developed for Speech Identification Scores, each of the subjects was tested for the following:

a) Establishment of SIS
b) Performance Intensity- Phonetically Balanced (PI-PB) function of the word lists.

a) Establishment of SIS

Each list (List 1 and List 2) was presented at intensity, 40 dB SL (reference: SRT). All the subjects were tested at this intensity level and each subject was tested in both the ears. The number of monosyllabic words correctly identified in each list was noted. The order of words in the lists was randomized in order to avoid the familiarity effect.

b) Performance Intensity- Phonetically Balanced (PI-PB) function of the word lists

The word identification of the PB word lists were established at different intensities, starting from 0 dBSL to 10 dBSL (reference: SRT) progressing in 2 dB steps. The subjects were instructed to repeat the test words and the responses noted down. It was ensured that all the subjects heard the different intensities. At each intensity, both lists (1&2) were presented. At each intensity level, the order of words in the lists was randomized in order to avoid familiarity effect. The average percentage correct scores for both the lists (1&2) were plotted as a function of intensity. This is called the Performance Intensity – Phonetically Balanced (PIPB) function.

3.9 Scoring of responses:

The responses were recorded on a score sheet for both SRT and SI tests by the tester. The number of correct responses was noted down for each of the lists.

3.10 Reliability check:

10 % of subjects were subjected to retesting for a time gap of at least five days. Test- retest reliability was calculated using this data.

3.11 Analysis:

Analysis was carried out to obtain the following information:

I. With respect to the SRT test:

   i. Correlation of SRT with PTA for frequencies 500 Hz, 1 KHz and 2 KHz.

   ii. Comparison of the SRTs obtained with the two lists (List I & II).

   iii. Effect of presentation level on the intelligibility of the SRT lists.

   iv. Presence/ absence of ear effect on the SRTs obtained.

II. With respect to the SI test:

   i. Comparison of difficulty level of the two SI lists (List 1 & 2).

   ii. Effect of presentation level on the intelligibility of the lists.

   iii. Presence/ absence of ear effect on the SI scores obtained.

3.12 Statistical analysis:

Appropriate statistical analyses were carried out for the data.

Chapter 4

RESULTS & DISCUSSIONS

The present study was carried out with an aim of developing and standardizing spondees and phonetically balanced word lists for speech recognition and identification tests in Mizo language. A total of one hundred (100) adults aged between 18 to 40 years and whose native language is Mizo participated in the study.

The following statistical analyses were done:

- Descriptive statistics of the scores obtained for the two lists of Spondee words and the two lists of monosyllabic words.
- Mixed Analysis of Variance to study the overall age, gender and list effects on the scores.
- Bonferroni Multiple Pairwise Comparison test for both spondees and PB word lists to check if any significant differences exist between different levels of presentation (i.e. 0 dBSL to 10 dBSL) for both.
- Paired sample t-test to compare list I & II of SRT and SI at each intensity level of presentation.

Spoondee word lists

4.1. Results of mean and standard deviation of SRT for the spondee word lists I & II

The mean and standard deviation of SRT for the Spondee word lists I & II for left and the right ear and the two genders are tabulated in table 1.

Table 1

Mean and Standard deviation (S.D) of SRT for spondee word lists I & II for Right ear and Left ear across gender.

	Spondee word list I				Spondee word list II			
	Male		Female		Male		Female	
	Left	Right	Left	Right	Left	Right	Left	Right
Mean (dBHL)	13.46	13.37	13.12	12.91	13.43	13.38	13.12	13.29
S.D (dBHL)	3.33	3.49	2.54	2.85	3.34	3.48	2.70	2.90

It is evident from Table 1 that the mean SRT scores for both the lists across both the genders and ears (left & right) are comparable. Figure-1 shows the graphical representation of the same results.

The mean SRT, considering both males & females and right & left ear together using list I was attained at 13.21 dB HL (re: PTA) with a SD of 3.05 and that for list II was attained at 13.30 dB HL (re: PTA) with a standard deviation (SD) of 3.10.

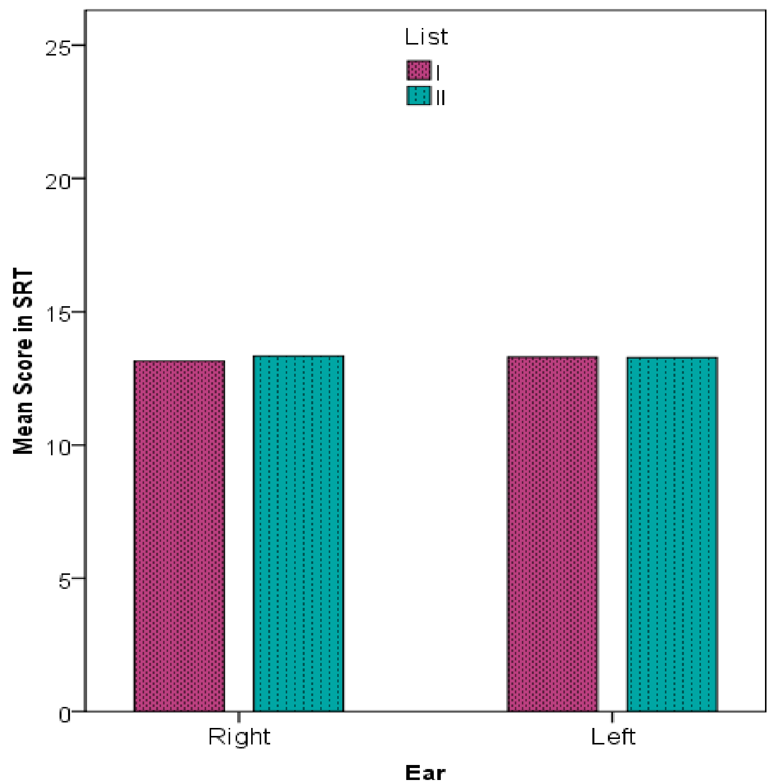

Figure 1: Mean SRT scores for right and left ear for List I & II.

Mixed analysis of Variance (ANOVA) was done to see if there is any statistical difference between the lists, between the ears, between genders. The results of Mixed ANOVA revealed that there was no significant difference between the lists [F (1, 98) = 0.726, p > 0.05], no significant difference between the ears [F (1, 98) = 0.049, p > 0.05] and no significant difference was found for gender [F (1, 98) = 0.264, p> 0.05]. Also, no significant interaction was found between the lists & gender [F (1, 98) = 0.948, p > 0.05], ear & gender [$F(1, 98) = 0.013$, $p > 0.05$], list & ear [$F(1, 98) = 1.080$, $p > 0.05$] and also for the three factor interactions i.e. list, ear and gender [$F(1, 98) = 0.619$, $p > 0.05$].

Thus, the averaged data of the hundred participants showed no significance in scores for spondee word lists I and II for right and left ear across gender at 0.05 confidence levels. Also, the results demonstrated that the two lists of spondaic words yield equivalent SRTs.

These findings are in agreement with that of the CID Auditory test W- 1 and W-2 spondee word lists developed by Hirsh et al., 1952, where the averaged data of several normal hearing subjects showed no consistent difference in the lists.

The findings of the present study are also in consonance with various Indian studies. Swarnalatha (1972) obtained SRT for spondees at 9 dBHL (re: PTA), Ghosh (1986) obtained SRT at 12 dBHL re: PTA), Tanuza (1984) obtained SRT at 13 dBHL for spondees in Manipuri language and Behera (2004) obtained SRT at 10 dBHL (re: PTA) for Oriya language.

4.2. Performance Intensity function of the Spondee list I & list II

Mean and SD of performance intensity function for spondee word lists I & II were calculated and are given in Table 2.

Table 2

Mean and SD of scores on spondee lists I & II at different intensity levels (0 to 10 dBSL with reference to pure tone average) across gender (raw scores).

	Intensity (dBSL) Re : PTA	**Female**		**Male**	
		Mean	**SD**	**Mean**	**SD**
Spondee List- II	Zero	10.85	2.27	11.30	2.37
	Two	16.10	2.83	15.86	2.82
	Four	20.50	2.35	20.32	2.57
	Six	23.37	1.59	23.65	1.54
	Eight	24.83	0.51	24.84	0.41
	Ten	25.00	0.00	25.00	0.00
Spondee List- I	Zero	11.47	1.90	11.67	2.12
	Two	16.66	2.54	16.53	2.99
	Four	20.85	2.28	21.00	2.48
	Six	23.68	1.51	23.98	1.37
	Eight	24.91	0.34	24.92	0.38
	Ten	25.00	.00	25.00	0.00

From table 2 we can see that the performance intensity function increases as the intensity is increased and almost reaches to a saturation level between 8 to 10 dBSL. It can also be seen that the scores do not differ across Males and Females. 50% correct criteria is met between 0 & 2 dB SL (ref: PTA). ANSI (1989) reported that, the 50% correct criteria for spondee words is obtained at 0 dBHL for normal hearing young adults. The finding of the present study is consistent with the correlation between SRT & PTA reported in the literature, thus validating the speech material developed.

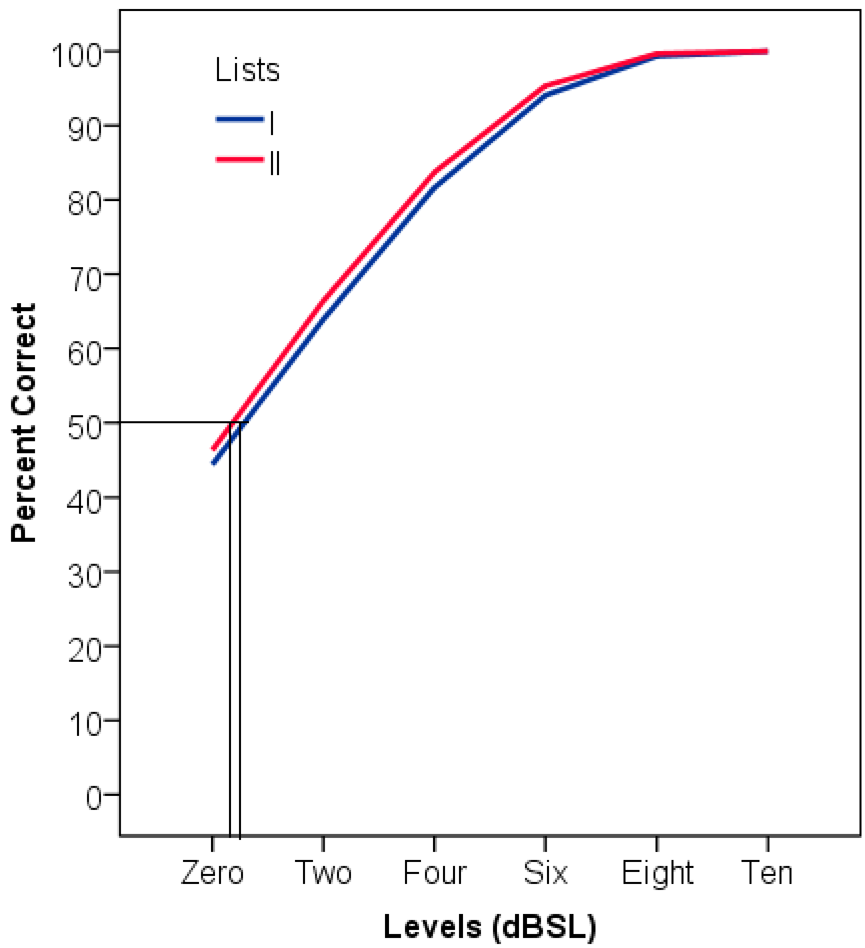

Figure 2. Performance intensity function of spondee Lists I and II across the six intensity levels.

Mixed ANOVA was done to test the statistical significance of the lists across the different intensity levels and also to see the interaction effects between different variables. Mixed ANOVA revealed that at 0 dBSL, 2 dBSL and 4 dBSL there was a statistical significance between the two spondee lists I & II [F (1, 98) = 8.592, p < 0.01], however, at higher presentation levels of 6 dBSL [t (99) = 1.976, p > 0.05] and 8 dBSL [t (99) = 1.469, p > 0.05] ,there was no significant difference between the two lists and the scores at lists I & II are equal at 10 dBSL . Also there was a statistical significance between list & level interaction [F (5, 490) = 3.588, p < 0.01]. However, there was no significant difference across genders [F (1, 98) = 0.080, p > 0.05, list & gender interaction [F (1, 98) = 0.017, p > 0.05], level & gender interaction [F (5, 490) = 0.502, p > 0.05] and interaction between lists, levels and genders [F (5, 490) = 0.502, p > 0.05].

The parameters which showed significant difference (lists and levels) were further analysed for pair wise comparison using the Bonferroni Multiple Comparison Test. Bonferroni's multiple comparison test revealed that the scores at all levels were significantly different from one another at 5% level of significance. List I across levels [F (5, 495) = 1557.594, p < 0.001 and list II across levels [F (5, 495) = 1556.283, p < 0.001].

This indicates that the intelligibility of the lists improved significantly with increase in presentation level. A study conducted by Hirsh et al. (1952) reported that, with increase in presentation levels, the identification scores for bisyllabic words increase.

This indicates that the intelligibility of the lists improved significantly with increase in presentation level. A study conducted by Hirsh et al. (1952) also reported that, with increase in presentation levels, the identification scores for bisyllabic words increases.

Mean SRT across List I and II across the six levels (figure 2) showed that as the intensity level was increased from 0 dBSL to +10 dBSL, the performance intensity function showed a steeply rising curve from 0 dBSL to 8 dBSL. However, at +10 dBSL, the curve flattens out as the intelligibility reached 100 %. Beyond this level, even with an increase in intensity, the scores would remain constant. This showed that, there was an increase in scores as the intensity of the presentation level is increased.

The performance intensity function for W-1 showed similar results where the intelligibility of the spondee words increases rapidly with increase in intensity. The articulation score risen from 0 to 100 % within a range of about 20 dB. There was an increase from 20 to 80% within a range of 8 dB and the scores reached the 100% point at about +14 dB above threshold.

Katz (2002) said that, in normal hearing individuals, the performance- intensity functions for spondaic words are very steep. The average level for 100% correct (maximum) score to first occur is approximately 7.5 dB HL. Similar findings have been reported by several authors have comparable results with the present study- Portuguese SRT materials (Harris et al., 2001), Mandarin Chinese SRT materials (Nissen et al., 2005), Arabic SRT materials (Ratcliff, 2006) etc.

Table 2 showed that as the presentation level was increased from 0 dBSL to 10 dBSL, the standard deviation of scores on spondee List I & II reduced from 2.27 to 0 and 1.90 to 0 respectively, indicating that at higher presentation levels, the variance was lesser.

In agreement with our findings, a study on ‘Development and evaluation of Mandarin disyllabic materials for Speech Audiometry in China by Wang et al (2007), Standard deviations reduced as the presentation level was increased (from 0 to 15 dB HL) in the mean performance-intensity function test. This indicates that at higher presentation levels the subjects' performance became less variable.

Monosyllabic Phonetically Balanced Word Lists

4.3. Results of mean and standard deviation of SIS for monosyllabic PB word lists I & II

The mean and standard deviation of SIS for the monosyllabic word lists I & II for left and the right ear and the two genders are tabulated in Table 3.

Table 3

Mean and Standard deviation (S.D) of SIS for PB word lists I & II for Right ear and Left ear across gender and across Ear.

	PB word list-1				PB word list-2			
	Male		Female		Male		Female	
	Left	Right	Left	Right	Left	Right	Left	Right
Mean(%)	97.52	99.26	97.24	97.2	97.88	97.52	97.28	96.82
S.D (%)	2.58	2.52	3.90	2.78	2.38	2.70	2.34	2.84

It is evident from the table that the mean SIS scores across both the genders are comparable. Also it can be seen that the mean scores for the SIS is almost similar across the ears i,e there is not much difference between the left ear and right ear scores. Thus, the two lists are equivalent. Figure 3 shows the graphical representation of the same results.

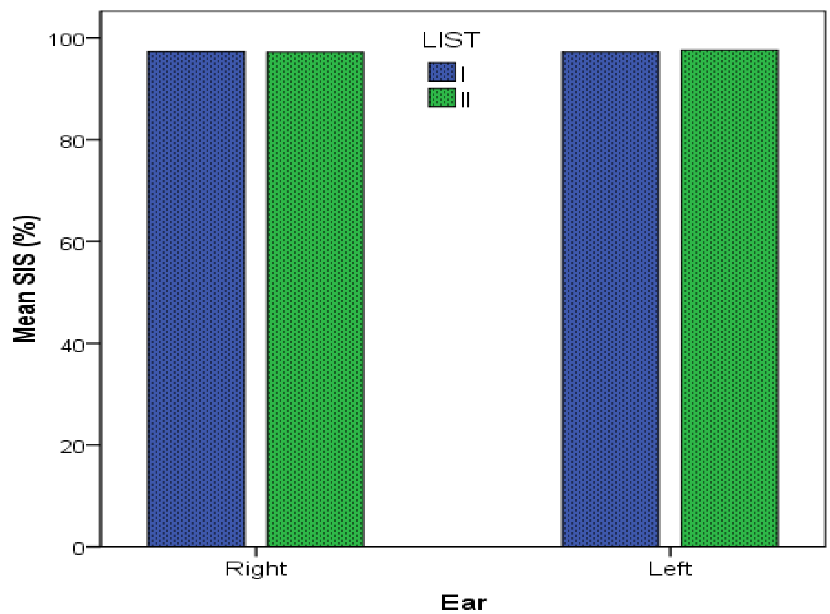

Figure 3. Mean SIS scores for Right and Left ear for PB List-1 and List-2.

Mixed analysis of Variance (ANOVA) was done to see if there is any statistical difference between the lists, between the ears, between genders. The results of Mixed ANOVA revealed that there was no significant difference between the lists [$F (1, 98) = 0.647$, $p > 0.05$], no significant difference between the ears [$F (1, 98) = 0.159$, $p > 0.05$] and no significant difference was found for gender [$F (1, 98) = 1.618$, $p > 0.05$]. Also, no significant interaction was found between the lists & gender [$F (1, 98) = 0.242$, $p > 0.05$], ear & gender [$F (1, 98) = 1.424$, $p > 0.05$], list & ear interaction [$F (1, 98) = 1.424$, p >

0.05] and also for the three factor interactions i.e. list, ear and gender [$F (1, 98) = 0.631$, $p > 0.05$].

This finding is in accord with that of the CID Auditory Test W 22, developed by Hirsh et al., 1952, where the monosyllabic word lists were divided into 4 lists, at 40 dB SPL, they found no significant differences between the speech identification scores.

Studies in other Indian languages have also yielded similar results, where they found no significance between their lists. Abrol (1971) obtained 100 % score in SIS at 30 dBSL (ref: PTA) for PB word lists developed in Hindi, Kapur (1971) obtained 100% score at 45 dB (ref:SRT) for Tamil PB word lists. The results are in accord with that of other studies done in Indian languages like Manipuri language, Bengali, Kannada and Oriya.

4.4. Performance Intensity function for Phonetically Balanced (PB) word list I & list II

Table 4

Mean and Standard deviation (S.D) of SI (in %) for PB word lists I & II at different intensity levels (0 to 10 dBSL with reference to SRT) across gender.

	Intensity (dBSL) (re: SRT) ↓	**Subjects (N = 100)**	
		Mean	**SD**
List-1	Zero	24.48	9.09
	Two	43.60	11.93
	Four	63.70	12.68
	Six	81.00	9.95
	Eight	93.70	5.49
	Ten	99.78	1.05
List-2	Zero	24.06	9.20
	Two	43.70	11.19
	Four	64.40	12.97
	Six	81.14	10.55
	Eight	94.04	6.18
	Ten	99.78	1.36

From table 4 we can see that mean raw scores of SIS increases as the intensity is increased and also the SIS does not differ across Males and Females.

For PB word lists I and II, the SD was more at the lower sensation level indicating greater dispersion of scores. The SD was lesser for higher sensation level reflecting lesser variance.

The results are in consonance with findings of Swarnalatha (1972), Mayadevi (1974), Tanuza (1984), Ghosh (1986) and Behera (2004).

As it can be seen from figure 4 that as the intensity level increases, the mean scores of SIS also increases and almost reaches the saturation level between 8 to 10 dBSL. The scores of SI are reaching to 100% at 10 dBSL.

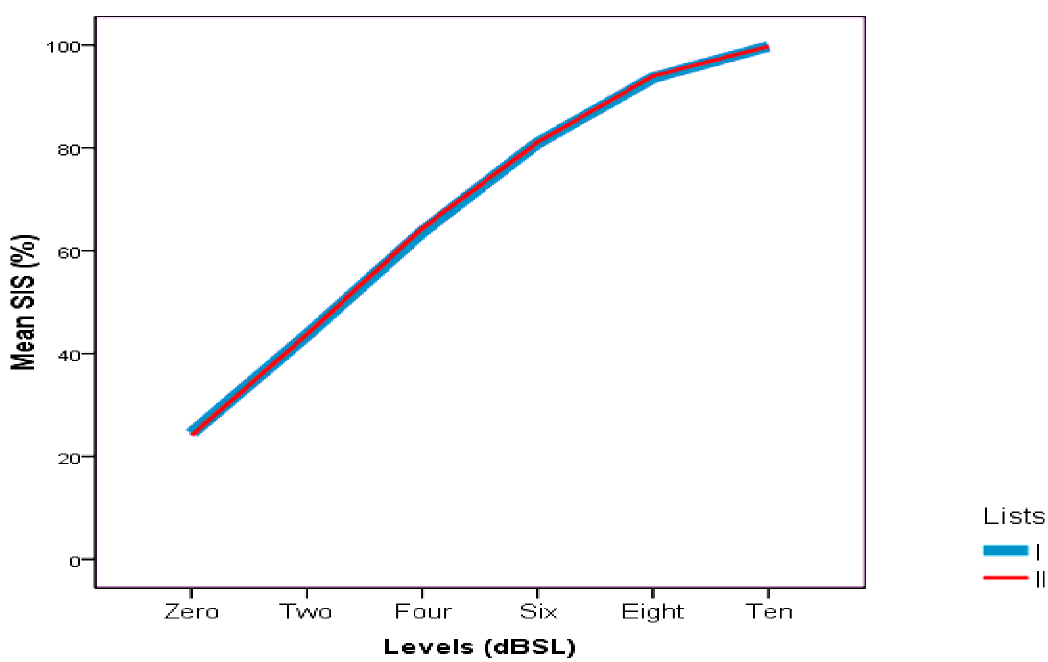

Figure 4. Mean SIS (in percentage) for PB Lists I and II across the six intensity levels.

Mixed ANOVA was done to see the statistical significance of the lists across the six intensity levels and also to see the interaction effects between different variables. Mixed ANOVA revealed that there was no statistical significance between the two PB word lists [$F (1, 98) = 0.026$, $p > 0.05$], list & level interaction [$F (5, 490) = 1.273$, $p > 0.05$], across genders [$F (1, 98) = 2.188$, $p > 0.05$, list & gender interaction [$F (1, 98) = 0.026$, $p > 0.05$], level & gender interaction [$F (5, 490) = 1.658$, $p > 0.05$] and interaction between lists, levels and genders [$F (5, 490) = 1.097$, $p > 0.05$. However, there was a statistical significance at the six intensity levels [$F (5, 490) = 2353.716$, $p < 0.01$].

The parameters which showed significant difference (lists and levels) were further analysed for pair wise comparison using the Bonferroni Multiple Comparison Test. Bonferroni's multiple comparison test revealed that all the six levels were significantly different from one another at 5% level of significance. List I across levels [$F (5, 495) = 1425.397$, $p < 0.001$ and list II across levels [$F (5, 495) = 2417.422$, $p < 0.001$].

Further Paired sample t- test was done to compare the two lists at different intensity levels. Paired sample t- test revealed that there is no significant difference between list I & II at all the levels: 0 dBSL [$t (99) =1.272$, $p > 0.05$], 2 dBSL [$t (99) = 0.069$, $p > 0.05$], 4 dBSL [$t (99) = 0.771$, $p > 0.05$], 6 dBSL [$t (99) = 0.148$, $p > 0.05$] and 8 dBSL [$t (99) = 0.552$, $p > 0.05$]. Equal scores were found at 10 dBSL.

From this study it can be seen that the mean scores for the SIS is almost similar across the gender and ears i.e. there is not much difference between the left ear and right ear scores at 0.05 confidence level. And the mean scores for the intelligibility of the PB word lists at increasing intensities showed that as the intensity level increases, the speech identification scores increases, and reached a score of 100% at 10 dBSL (with reference to SRT) for both the lists. The curve is sharply rising indicating a positive relationship with the percentage scores and the level at which the material is presented. The maximum score on the PI-PB function is called the PB max.

These findings are in high correlation with that of earlier studies done by several authors where essentially a same curve is obtained for normal hearing subjects. Maroonroge and Diefendorf (1984) in their study of speech identification scores for 3 word lists- NU-6, California Consonant Test and Pascoe high frequency word lists found that the speech identification scores tend to improve up to about 30 dBSL for normals.

For the CID Auditory Test W-22, it was found that the scores increased sharply with increasing levels of presentation and scores remained constant at about 40 dBSPL. (Hirsh et al., 1952).

Hood & Poole (1980), based on their study had stated that, in normal hearing subjects, the performance intensity function (curve) derived for a single list is no different from that obtained with a number of equivalent lists. Lau and So (1988), on their study of Cantonese Speech Audiometry where they found Speech discrimination scores at various stimulus levels for 10 monosyllabic word lists, they stated that, averaged discrimination scores over all lists tends to increase and the Standard deviation decreases as the level increases, they found that at 30 dBHL all the 10 lists have equivalent intelligibility.

Similar findings have been listed on studies relating to Speech Identification for monosyllabic words across different languages.

4.5. Reliability:

Reliability check was performed on 10% of the obtained data. Cronbach's Alpha Coefficient was done to check the reliability of the data.

Table 5

Percent reliability of Spondee list I & II and PB list 1 & 2.

Lists	**SRT**	**SIS**
List I	95%	83%
List II	89%	85.5%

Table 5 shows Cronbach's Alpha Coefficient thus indicated that there is correlation for SRT and SIS of lists I and II. Thus this shows that the Spondee word lists and the Phonetically Balanced word lists are reliable.

Chapter 5

Summary and Conclusion

Speech audiometry tests are essential components of any comprehensive audiological evaluation. The hearing impairment inferred from a pure tone audiogram cannot depict, beyond the gross generalizations, the degree of disability in speech communication caused by a hearing loss. It is logical that tests of hearing function should be performed with speech stimuli. Using speech audiometry, audiologists set out to answer questions regarding patients' degree of hearing loss for speech, the levels required for their most comfortable levels and uncomfortable loudness levels, the range of comfortable loudness and perhaps most importantly, their ability to recognize the sounds of speech. Speech-language pathologists and audiologists use reported findings of speech audiometric results useful in both therapy planning and counseling (Martin & Clark, 2003).

The purpose of this study was to develop and standardize Spondees and PB words in Mizo language that can be used to measure the Speech Recognition Threshold (SRT) and Speech Identification Scores (SIS) for native speakers of Mizo, a language spoken in Mizoram, India. Two lists each of Spondees and PB words with high familiarity were developed as per standard procedures and the SRT and SIS evaluated for 100 (one hundred) native speakers of Mizo language in the age ranged between 18 years to 40 years, with normal hearing. Performance – intensity function for each word lists were evaluated at 6 intensity levels (0 to 10 dBSL) in 2 dB increments. In order to establish the reliability of these materials, 10% of the subjects were retested after a minimum period of 5 days.

The results of the study revealed:

- No significant difference in scores between the two Spondee lists and PB word lists for both the right and the left ear across gender.
- The two lists of spondaic words and PB words yielded equivalent SRTs and SIS (at 40 dB SL, ref: SRT).
- At six different levels of presentation, there was a significant difference in scores between spondee lists I & II at lower presentation levels (0 dBSL, 2 dBSL and 4 dSL). However, as the presentation level increase to 6 dBSL and 8 dBSL, no significant difference in scores were found and at 10 dBSL of presentation, the scores of list I & II were equal.
- As the intensity increases, the scores were found to increase.
- There was no significant difference in scores between PB word lists 1 & 2 at six different levels of presentation (0 dBSL to 10 dBSL).
- As the intensity increases, the scores were found to increase.
- For Spondees, the scores almost reached 100% at about 8 dBSL (reference PTA) and for PB words, the scores reached 100% at 10 dBSL (reference SRT).

The results of the study were in accord with previous studies done in different languages. The materials were found to have excellent reliability. Thus the Spondee and the PB word lists developed can be used in clinical situations for Speech Audiometry.

5.1 Limitations of the study:

Due to unavailability of phonetically balanced word lists in Mizo language, attempts were made to phonetically balance the word lists by inferring information regarding the same with language samples from books, newspapers and telephone conversations. Frequency of occurrence of phonemes derived at by this method needs to be verified.

5.2 Future research directions:

- There have been no materials developed for speech audiometry for children in Mizo language. So, extensive studies could be carried out to develop speech audiometry materials for different age groups.
- These tests could be used to evaluate the speech perception abilities through different hearing aids and thus, its utility in hearing aid selection can be assessed.
- The tests can be used to assess the utility of different devices such as frequency modulation (FM) systems, cochlear implants etc.
- It can also be used to assess the efficacy of intervention with different therapy programs.

REFERENCES

American National Standards Institute (1989). *American National Standards for audiometers*. (ANSI S3. 6- 1989). New York: American National Standards Institute.

American National Standards Institute (1991). *American National Standards for maximum permissible ambient noise levels for audiometric test rooms*. (ANSI S3. 1- 1991). New York: American National Standards Institute.

American National Standards Institute (1996). *American National Standards specification for audiometers*. (ANSI S3.6- 1996). New York: American National Standards Institute.

American Speech- Language- Hearing Association (1979). Guidelines for determining the threshold level for speech. *ASHA, 21*, 353-356.

American Speech- Language- Hearing Association (1988). Guidelines for determining the threshold level for speech. *ASHA, 30*, 85-89.

American Speech- Language- Hearing Association (1998). Guidelines for determining threshold level for speech. *ASHA*, *30*, 85- 89.

Aspinall, K. B. (1973). *The effect of phonetic balance on discrimination for speech in subjects with sensorineural hearing loss*. Unpublished Doctoral dissertation. University of Colorado.

Beattie, R. C., Forrester, P. W., & Ruby, B. K. (1976). Reliability of the Tillman- Olsen procedure for determination of spondee threshold using recorded and live voice presentations. *Journal of the American Audiological Society*, *2*, 159-162.

Behera, S. (2004). *Development and standardization of speech perception test in Oriya.* Unpublished Master's dissertation. University of Mysore, India.

Bell, T. S., & Wilson, R. H. (2001). Sentence recognition material based on frequency of word use and lexical confusability. *Journal of the American Academy of Audiology*, *12*, 514-522.

Bess, F. H. (1983). Clinical assessment of speech recognition. In D. F. Konkle & W. F. Rintelmann (Eds.). *Principles of Speech Audiometry*. (127- 201). Baltimore: University Park Press.

Bilger, R. C. (1984). Speech recognition test development. In E. Elkins, (Ed). Speech recognition testing by the hearing impaired. *ASHA Reports*, *14*, 2- 15.

Black, J. W. (1957). Multiple choice Intelligibility tests. *Journal of Speech and Hearing Disorders*, *22*, 213- 235.

Bowling, L., & Elpern, B. (1961). Relative intelligibility of items on CID auditory test W-1. *Journal of Audiitory Research*, 1, 152-57.

Boothroyd, A. (1968). Developments in speech audiometry. *Sound*, *2*, 3-10.

Carhart, R. (1946). Monitored live voice as a test of auditory acuity. *Journal of the Acoustical Society of America, 17*, 339-349.

Carhart, R. (1946). Speech reception in relation to pattern of pure tone loss. *Journal of Speech Hearing Disorder*, *11,* 97- 108.

Carhart, R. (1950). Hearing aid selection by university clinics. *Journal of Speech & Hearing Research*, *15,* 103- 113.

Carhart, R. (1971). Discussion, questions, answers, comments. In Rosjskjer, C. (Ed.). *Speech Audiometry*. (229). Denmark: Second Danavox Symposium.

Carhart, R. & Jerger, J. (1959)**.** Preferred method for clinical determination of pure tone thresholds. *Journal of Speech and Hearing Disorders,* 24, 330.

Causey, G. D., Hood, L. J., Hermanson, C. L., & Bowling, L. S. (1984). The Maryland CNC Test: Normative Studies. *Audiology, 23,* 552- 568.

Chaiklin, J. B. & Ventry, I. M. (1964). Spondee threshold measurement: a comparison of 2 and 5 dB steps. *Journal of Speech and Hearing Disorders*, 29, 47- 59.

Clemis, J. & Carver, W. (1967). Discrimination scores for speech in Meniere's disease. *Archieves of Otolaryngology*, *86,* 614-618.

Cowan, R. S., Deldot, J., Baeker, C. J., Sarant, J. Z., Pegg, P., et al. (1997). Speech perception results for children with implants with different levels of preoperative residual hearing. *American Journal of Otology*, 18, 5125- 5126.

Creston, J. E., Gillespie, M., & Krahn, C. (1966). Speech Audiometry: Taped vs. live voice. *Achieves of Otolaryngology*, *83*, 14-17.

Davis, H. & Silverman, S. R. (1960). *Hearing and Deafness*. Rev. ed. Holt, Rinehart & Winston, New York.

Davis, H. & Silverman, S. R. (1978). *Hearing and Deafness*. (4^{th} ed.), Holt, Rinehart & Winston: New York.

De, N. S. (1973). Hindi PB list for speech audiometry and discrimination test. *Indian Journal of Otolaryngology, 25,* 64-75.

Dirks, D. D. & Morgan, D. E. (1983). Measures of discomfort and most uncomfortable loudness. In Konkle, D. F. & Rintelmann, W. F. (Ed.). *Principles of Speech Audiometry* (203- 229). Baltimore: University Park Press.

Dowell, R. C., Mecklenburg, D. J., & Clark, G. M. (1986). Speech recognition for 40 patients receiving multi- channel cochlear implants. *Archives of Otolaryngology, Head and Neck Surgery, 112*, 1054- 1059.

Egan, J. P. (1948). Articulation testing methods. *Laryngoscope*, 58, 955-981.

Egan, J. J. (1979). Basic aspects of speech audiometry. *Ear, nose, throat Journal*, *58*, 190- 193.

Elliot, L. L., & Katz, D. (1980). *Development of a new children's test of speech discrimination.* St. Louis, MO: Auditec.

Fairbanks, G. (1958). Test of phonemic differentiation: The Rhyme Test. *Journal of the Acoustical Society of America*, *30,* 596-600.

Fletcher, H. & Steinberg, J. C. (1929). Articulation testing methods. *Bell System Technology Journal*, 7, 806- 854.

Fletcher, H. (1950). A method for calculating hearing loss for speech from an audiogram. *Journal of the Acoustical Society of America*, *22*, 1- 5.

Fletcher, H. (1965). *Speech, Hearing and Communication*. Van Nostrand: New York.

Fulton, R. T. & Llyod, L. L. (1975). *Auditory Assessment of the Difficult- to- test.* Baltimore: William & Wilkins.

Gardner, H. J. (1971). Application of high frequency consonant discrimination word test in hearing aid evaluation. *Journal of Speech and Hearing Disorders*, *36,* 354- 355.

Gat, I. B. & Keith, R. W. (1978). An effect of linguistic experience: Auditory discrimination by native and non- native speakers of English. *Audiology, 17,* 339- 345.

Gelfand, S. A. (1975). Use of the carrier phrase in live- voice speech discrimination testing. *Journal of Audiological Research*, *15,* 107- 110.

Gelfand, S. A. (1993). *A clinical speech recognition method to optimize reliability and efficiency*. Paper presented at Convention of American Academy of Audiology, Phoenix.

Gelfand, S. A. (1998). Optimizing the reliability of speech recognition scores. *Journal of Speech, Language and Hearing Research*, *41*, 1088- 1102.

Gelfand, S. A. & Silman, S. (1985). Functional hearing loss and its relationship to resolved hearing levels. *Ear and Hearing*, *6*, 151- 158.

Gelfand, S. A. & Silman, S. (1993). Relationship of exaggerated and resolved hearing levels in unilateral functional hearing loss. *British Journal of Audiology, 27*, 29- 34.

Gelfand, S. A. (2001). *Essentials of Audiology*. Thieme: New York.

Ghosh, D. (1986). *Development and Standardization of speech materials in Bengali language.* Unpublished Master's dissertation. University of Mysore, India.

Gladstone, V. S. & Siegenthaler, B. M. (1971). Carrier phrase and speech intelligibility score. *Journal of Audiological Research*, *11*, 101-103.

Glaser, R. G. (1974). *Hearing aid evaluations using spectral density classified word list.* Ph. D. dissertation, Kent State University, Ohio.

Goetzinger, C. P. (1972). Word discrimination testing. In J. Katz (Eds.) (2nd Edn). *Handbook of Clinical Audiology*, Baltimore: William and Wilkins.

Hagerman, B. (1976). Reliability in the determination of speech discrimination. *Scandinavian Audiology*, *5*, 219-228.

Hagerman, B. (1979). Reliability in the determination of speech reception thresholds. *Scandinavian Audiology*, *8*, 195-202.

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scandinavian Audiology*, *11*, 79- 87.

Hagerman, B. (1984). Some aspects of methodology in speech audiometry, *Scandinavian Audiology Supplement*, *21*, 1-25.

Harris, R. W., Goffi, M. V. S., Pedalini, M. E. B., Gygi, M. A., & Merrill, A. (2001). Psychometrically equivalent Brazilian Portuguese trisyllabic words spoken by male and female talkers. *Pró-Fono, 13*(1), 37-53.

Haskins, H. L. (1949). *A phonetically balanced test of speech discrimination for children*. Unpublished Master's thesis, Northwestern University, Illinois.

Hirsh, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eldert, E., & Benson, R. W. (1952). Development of materials for speech audiometry. *Journal of speech and hearing disorders, 17*, 321-37.

Hood, J. D., & Poole, J. P. (1980). Influence of speaker and other factors affecting speech intelligibility. *International Journal of Audiology,* 19, 434- 455.

House, A. S., William, C. A., Hecker, M. H. L., & Kryter, K. D. (1965). Articulation-testing methods; consonantal differentiation with a closed response set. *Journal of the Acoustical Society of America, 37*, 158-166.

Hudgins, C. V., Hawkins, J. F., Karlin, J. E., & Stevens, S. S. (1947). Development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope*, *57*, 57-89.

Huff, S. J. & Nerbonne, M. A. (1982). Comparison of the American- Speech- Language-Hearing Association and the revised Tillman- Olsen methods for speech threshold measurement. *Ear and hearing*, *3*, 335- 339.

Jahner, J. A., Schlauch, R. A., Doyle, T. (1994). A comparison of American Speech-Language- Hearing Association guidelines for obtaining speech- recognition thresholds. *Ear and Hearing*, *15*, 324- 329.

Jerger, J., Speaks, C., & Tramell, J. (1968). A new approach to speech audiometry. *Journal of Speech and Hearing Disorders, 33*, 318- 328.

Jerger, S., Jerger, J., & Abrams, S. (1983). Speech audiometry in the young child. *Ear and Hearing*, *4*, 56-66.

Johnson, E. W. (1977). Auditory test results in 500 cases of acoustic neuroma. *Archives of Otolaryngology*, *103*, 152-158.

Kalikow, D. N., Stevens, K. N., & Elliot, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with connected word predictability. *Journal of the Acoustical Society of America*, *61*, 1137- 1351.

Kapur, Y. P. (1971). *Needs of the speech and hearing handicapped in India*. Christian Medical College and Hospital, Vellore.

Katz, J. (2002) (Ed.), *Handbook of Clinical Audiology*. Baltimore: Lippincott Williams & Wilkns.

Lalnghinglova (2003). *Nitin Lalpa ke bulah*. Lengchhawn press: Aizawl.

Lau, C. C., & So, K. W. (1988). Material for Cantonese speech audiometry constructed by appropriate phonetic principles. *British Journal of Audiology, 22*(4), 297-304.

Lehiste, I. & Paterson, G. E. (1959). Linguistic considerations in the study of speech intelligibility. *Journal of the Acoustical Society of America, 31*, 280- 286.

Lerman, J. W., Ross, M., & McLaughlin, R. M. (1965). A picture-identification test for hearing- impaired children. *Journal of Auditory Research*, *5*, 273-278.

Mallikarjuna (1990). Spondee words and Phonetically Balanced Monosyllabic words in Gujarati language. In S.K. Kacker & V. Basavaraj (Ed.). *Indian Speech Language Hearing Tests*- The ISHA Battery. AIISH, Mysore.

Maroonroge, S. & Diefendorf, A. O. (1984). Comparing normal hearing and hearing-impaired subject's performance on the Northwestern Auditory Test Number 6, California Consonant Test, and Pascoe's High-Frequency Word Test. *Ear and Hearing*, 5(6), 356-60.

Martin, A. (2001). Hearing in elderly: A population study. *Audiology*, *40*, 285- 293.

Martin, F. N., & Clark, J. G. (2003). *Introduction to Audiology* (8th Edition). New Jersy: Prentice Hall, Engelwood Cliffs.

Martin, F. N., Hawkins, R. R., & Bailey, H. A. (1962). The non- essentiality of the carrier phrase in phonetically balanced (PB) word testing. *Journal of Audiological Research*, *2*, 319-322.

Martin, F. N., & Stauffer, M. L. (1975). A modification of the Tillman-Olsen method for obtaining the speech reception threshold. *Journal of Speech and hearing Disorders*, *40*, 25-27.

Mascarenhas, K. E. (2002). *A high frequency Kannada Speech Identification Test (HF-KSIT).* Unpublished Master's dissertation. University of Mysore, India.

McLennan, R. O., & Knox, A. W. (1975). Patient- controlled delivery of monosyllabic words in a test of auditory discrimination. *Journal of Speech and Hearing Disorders*, *40*, 25-28.

Mayadevi (1974). *Development and Standardization of Common Speech Discrimination Test for Indians.* Unpublished Master's dissertation. University of Mysore, India.

Nagaraja, M. N. (1973). *Development of Synthetic Speech Identification Test in Kannada Language.* Unpublished Master's dissertation. University of Mysore, India.

Nagaraja, M. N. (1990). Testing, interpreting and reporting procedures in speech audiometric tests. In S.K. Kacker and V. Basavaraj (Ed.), *Indian Speech Language Hearing Test*- The ISHA Battery.

Newby, H. A. (1972). *Audiology* (2nd Ed.). Appleton- New York: Century- Crofts.

Nillsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise test for the measurement of speech reception threshold in quiet and in noise. *Journal of the Acoustical Society of America*, *95*, 1085- 1099.

Nissen, S. L., Harris, R. W., Jennings, L., Eggett, D. L., & Buck, H. (2005). Psychometrically equivalent Mandarin bisyllabic speech discrimination materials spoken by male and female talkers. *International Journal of Audiology, 44*(7), 379-390.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition of children and older adults. *Journal of the Acoustical Society of America*, *87,* 2705-2715.

Olsen, W. O., Noffsinger, D., & Kurdziel, S. (1975). Speech discrimination in quiet and in white noise by patients with peripheral and central lesions. *Archives of Otolaryngology, 96,* 231- 247.

Olsen, W. O., & Matkin, N. D. (1979). Speech Audiometry. In W. F. Rintelmen. (Ed.). *Hearing Assessment*. Baltimore: University of Park Press.

Olsen, W. O., Van Tasell, D. J., & Speaks, C. E. (1997). Phoneme and word recognition for words in isolation and in sentences. *Ear and Hearing*, *18*, 175- 188.

O'Neill, J. J., & Oyer, H. J. (1966). *Applied Audiometry*. New York: Dodd, Mead & Company, Inc.

Owens, E., & Schubert, E. D. (1972). Development of California Consonant test. *Journal of Speech and Hearing Research*, *20*, 463- 474.

Penrod, J. P. (1994). Speech threshold and word recognition/discrimination testing. In: Katz, J ed. *Handbook of Clinical Audiology*. 4th ed. (147- 164). Baltimore: William & Wilkins.

Plomp, R. (1986). A signal- to- noise ratio model for the speech perception threshold of the hearing impaired. *Journal of Speech and Hearing Research*, *29*, 146- 154.

Raffin, M. J. M. & Thornton, A. (1980). Confidence levels for differences between speech discrimination scores. *Journal of Speech and Hearing Research*, *23*, 5- 18.

Ramachandra, P. (2001). *High Frequency Speech Identification Test for Hindi and Urdu speakers*. Unpublished Master's dissertation. University of Mysore, India.

Ramkissoon, I. (2001). Speech recognition thresholds for multilingual populations. *Communication Disorders Quarterly*, 22 (3), 158-162

Rajashekhar, B. (1976). *The development and standardization of a picture SRT test for adults and children in Kannada*. Unpublished Master's dissertation. University of Mysore, India.

Ratcliff, E. R. (2006). *Psychometrically Equivalent Bisyllabic Words for Speech Reception Threshold Testing in Arabic*. Unpublished Master's thesis. Brigham Young University, Utah.

Resnick, D. M., & Becker, M. H. (1963). Hearing aid evaluation- A new approach. *American Speech and Hearing Association*, *5*, 695-699.

Ross, M., & Lerman, J. (1970). A picture identification test for hearing impaired children. *Journal of Speech and Hearing Research*, *13*, 44-53.

Roeser, R. J., Valente, M., Hosford-Dunn, H. (2000). *Audiology Diagnosis*. Thieme Medical publishers. Inc.

Samuel, J. D. (1976). *Development and standardization of phonetically balanced material in Tamil.* Unpublished Master's dissertation. University of Mysore, India.

Sarant, I. Z., Balmey, P. I., Dowel, R. C., Clark, G. M., & Gibson, W. P. (2001). Variation in speech perception scores among children with Cochlear Implants. *Ear and Hearing*, *22*, 18-28.

Silman, S., & Silverman, C. A. (1991). *Auditory Diagnosis: Principles and Applications*. San Diego: Academic Press.

Stach, B. A. (1998). *Clinical Audiology: An Introduction.* San Diego: Singular Publishing Group.

Swarnalatha, C. K. (1972). *Development and standardization of speech test material in English for Indians.* Unpublished Master's dissertation. University of Mysore, India.

Tanuza, E. D. (1984). *Development and Standardization of Speech test material in Manipuri language.* Unpublished Master's dissertation. University of Mysore, India.

Tillman, T. W., & Carhart, R. (1966). *An expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words*. Northwestern University Auditory Test No. 6. Technical report SAM- TR- 66-65. Brooks A. F. B., TX:USAF School of Aerospace Medicine.

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

Ullrich, K., & Grimm, D. (1976). Most comfortable listening level presentation versus maximum discrimination for word discrimination material. *Audiology*, *15*, 338- 347.

Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *Journal of the Acoustical Society of America*, *107 (3)*, 1671- 1684.

Wang, S., Mannell, R., Newall, P., Zhang, H., & Han, D. (2007). Development and evaluation of Mandarin disyllabic materials for speech audiometry in China. *International Journal of Audiology*, 46, 719- 731.

Weisleder, P., & Hodgson, W. R. (1989). Evaluation of Four Spanish Word-Recognition-Ability Lists. *Ear & Hearing,* 10(6):387-393.

Wilson, R. H. & Margolis, R. H. (1983). Measurements of Auditory Thresholds for Speech Stimuli. In Rintelmann, W. F. and Konkle, D. F. (Ed.) *Principles of Speech Audiometry*, Baltimore: University Park Press, 79-126.

Wilson, W. I., & Moodley, S. (2000). Use of the CID W22 as a South African English speech discrimination test. *South African Journal of Communication Disorders, 47,* 57-62.

Young, L. L., Dudley, B., & Gunter, M. B. (1982). Thresholds and psychometric function of the individual spondaic words. *Journal of Speech and Hearing Research, 25,* 586- 593.

APPENDIX I

Spondee word list I

1) Artui
2) Nauban
3) Buhfai
4) Lungkham
5) Banbun
6) Biakbuk
7) Dawhkan
8) Tukverh
9) Serthlum
10) Leilung
11) Thingrem
12) Vaivut
13) Zungbun
14) Vawmpuah
15) Kawrlum
16) Zikno
17) Chanchin
18) Chhangthawp
19) Hamrik
20) Hmuhnawm
21) Sahuan
22) Zanlai
23) Kehmawr
24) Hremhmun
25) Hnatlang

APPENDIX II

Spondee word list II

1) Arawn
2) Bakkilh
3) Biakin
4) Dumpawl
5) Ennawm
6) Kohhran
7) Tarmit
8) Thilpek
9) Sawmhnih
10) Rannung
11) Vawnban
12) Balhla
13) Bengbeh
14) Awngphah
15) Banrek
16) Bawkkhup
17) Ruhno
18) Chawhma
19) Fiamthu
20) Kalkawng
21) Sahriak
22) Zaizir
23) Lukhum
24) Chhungkhat
25) Hnathawh

APPENDIX III

PB word list I

1) Chaw
2) Cher
3) Ding
4) Nel
5) Hmui
6) Thap
7) Dur
8) Fur
9) Nghing
10) Hmun
11) Khar
12) Lung
13) Kum
14) Kut
15) Sil
16) Dah
17) Dum
18) Fim
19) Hil
20) Keh
21) Meng
22) Phah
23) Ring
24) Vur
25) Zeng
26) Hlawk
27) Ram
28) Par
29) Chaw
30) Thar
31) Chhawng
32) Dik
33) Kher
34) Ping
35) Thim
36) Chem
37) Hring
38) Zar
39) Tap
40) Vak
41) Sam
42) Bang
43) Zan
44) Hmel
45) Khum
46) Hrilh
47) Chhang
48) Bang
49) Hnar
50) Thap

APPENDIX IV

PB word list II

1) Khap
2) Hnung
3) Hling
4) Nal
5) Ser
6) Par
7) Hming
8) Hah
9) Lei
10) Mei
11) Sang
12) Sen
13) Rit
14) Ram
15) Phun
16) Zim
17) Buh
18) Chhun
19) Dar
20) Phum
21) Fing
22) Ngun
23) Zung
24) Kang
25) Chawk
26) Keng
27) Thawm
28) Khur
29) Pem
30) Thing
31) Var
32) Tlar
33) Chak
34) Far
35) Nghak
36) Tlang
37) Rul
38) Thlam
39) Chep
40) Dek
41) Chhin
42) Sin
43) Beng
44) Zum
45) Veng
46) Hmin
47) Dan
48) Khuh
49) Chhum
50) Hnam