DEVELOPMENT OF LOW FREQUENCY WORD LISTS IN HINDI AND KANNADA

Project funded by A.I.I.S.H Research Fund (ARF)

(2012-2013)

Sanction Number: SH/AUD/ARF/4.53/2012-2013

Total Grants: ₹ 5, 92, 000

Total Duration of the Project: 10.09.2012 – 08.11.2013

Principal Investigator
Prof. Animesh Barman
Professor of Audiology
Department of Audiology

Co-Investigators
Mr. Prashanth Prabhu
Mr. Niraj Kumar Singh
Dr. Vijay Kumar Narne
Department of Audiology

Research Officers
Ms. Mahima Gupta
Ms. Spoorthi T
Department of Audiology

All India Institute of Speech and Hearing
Manasagangothri, Mysore – 570 006

ACKNOWLEDGEMENTS

Our sincere gratitude to our Director, Prof. S R Savithri, All India Institute of Speech and Hearing, Mysore, for funding and providing the infrastructure to carry out the project.

We extend our heartfelt gratitude to Mr. Sharath Kumar, Research Assistant, department of Audiology. Our thanks are due to all the clinical staff and faculty of the department who have helped us in our data collection. We are also grateful to all the participants of the study and their caregivers for their valuable contribution.

Prof. Animesh Barman
Principal Investigator

Mr. Prashanth Prabhu
Co-Investigator

Mr. Niraj Kumar Singh
Co-Investigator

Dr. Vijay Kumar Narne
Co-Investigator

Ms. Mahima Gupta
Research Officer

Ms. Spoorthi T
Research Officer

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Conventional belief that only high frequency information is important for speech understanding is only partly true as recent studies have demonstrated the importance of low frequency information. However, not many standardized tests have been developed to assess the same. Thus, this research was taken up to develop, standardize and validate low frequency bi-syllabic word lists in two Indian languages viz. Hindi and Kannada.

During first phase, bi-syllabic words were collected, evaluated by native linguist and words were short listed based on familiarity ratings given by native speakers. Those words were then recorded 5 times each by a native male speaker. Best recorded words were selected through subjective and objective analysis. Further, energy of each word above and below 1.5 kHz was calculated using FFT and amplitude ratios were obtained. Later, using k-mean clustering, words with more energy (by around 20dB) below 1.5 kHz were separated from rest of the words. Equally difficult word lists were generated from these low frequency words by obtaining psychometric function curves and by calculating mean sensation level at which 50% SI scores were obtained and slope of the psychometric functions. As a result, a total of 10 and 7 word lists each were developed in Hindi and Kannada respectively.

In the next phase, lists were standardized by administering on 100 normal hearing Kannada and Hindi speakers. In addition, word lists were administered on 40 adult normal hearing participants at 0, 10, 20 and 30dB SLs. Results showed similar trend of increase in SI scores with increase in SL and reaching a maximum score (>90%) at around 30dB and 20dB SL in Hindi and Kannada. During the final phase, developed lists were validated on 10 low frequency cochlear hearing loss Kannada participants. Due to unavailability of native Hindi patient group, hearing loss was simulated using MATLAB and NIOSH hearing loss simulator software. Results of validation showed that low frequency word lists are sensitive enough to tap the speech understanding difficulty in the clinical population.

INTRODUCTION

Hearing loss can emerge due to a variety of causes. Auditory signs and symptoms of different hearing disorders are almost unique. One of the signs that help us in speculating the cause of the hearing problem is the shape of the audiogram. The configuration of loss that is reflected on the audiogram is usually different for different disorders. Certain disorders lead to flat pattern while few others result in more loss in either low or high frequencies. Presbycusis, Noise Induced hearing loss, Ototoxicity are a few of the disorders known to cause high frequency hearing loss while others like Auditory Dyssynchrony, Menieres Disease and Otoslerosis have been identified to usually lead to low frequency hearing loss (Wang et al., 2002; Harner, Fabry & Beatty, 2000; Knox & Mcpherson, 1997; Hannley, 1993; Tarter & Robins, 1990).

Many studies have been carried out to estimate the prevalence of different configurations of hearing loss over years. Margolis and Saly (2008) have reported a US based study which mainly analysed the databases of an academic health centre audiology clinic. Findings revealed that sloping configuration was the most common (40%), followed by flat (16%), peaked (5%), rising (3%), other (2%) and trough (1%) configurations. As per the findings of National Speech and Hearing Survey, carried out on 38568 children between grade 1 to 12, around 2% had low frequency hearing loss and 3 to 4% had high frequency hearing loss (Hull, Mielke, Timmons & Willeford, 1971). Rabinowitz, Slade, Galusha, Dixon-Ernst, & Cullen (2006) reported that among 2526 young adults between the age range of 17 to 25 years, had around 16% and 5 % of high and low frequency hearing losses respectively. In elderly individuals (Age >60 years), the prevalence of different hearing loss configuration was found to be 29% flat, 6% rising, 36% gradually sloping and 29% sharply sloping audiogram (Gates, Couropmitree & Myers, 1999). On analyzing these study results, it can observed that though the prevalence of low frequency loss is lesser compared to high frequency loss, across ages there is a small but considerable proportion of population being subjected to effects of low frequency loss.

As the principle avenue of the human communication and interaction, speech has been widely accepted as the most important signal that humans hear and use. It is a complex signal that spreads over a wide range of frequencies, roughly between 250 and 8000 Hz (Dobie & Van Hemel, 2004). The hearing loss in any part of this frequency range can impact the hearing and there by speech understanding. Conventionally, it was believed that high frequency sounds are more important for speech understanding and thus, more focus has been given to perception of speech by individuals with high frequency hearing loss. There is a lot of attempt done to develop assessment tools for testing high frequency speech perception difficulties (Eric, Benedicte, Jean-Fracois, Samia & Lionel, 2008; Sudipta & Yathiraj, 2006; Stelmachowicz, Pittman, Hoover, Lewis & Moeller, 2004; Nobel, Sinclair & Byrne, 1998; Pascoe, 1975; Gardner, 1971) and research focus has even extended towards developing new concepts and technologies for better rehabilitation of individuals with high frequency loss. However, the concept that only high frequency sounds help in better speech understanding is just partly true as there are researches supporting the importance of low frequencies in speech perception.

Jin and Nelson (2010) studied the effect of low to mid frequency information on sentence recognition and found that those listeners with more hearing losses in the low frequencies were poorest at understanding interrupted sentences. Also, low to mid frequency hearing thresholds accounted for most of the variability in masking release for listeners with hearing impairment. Based on these findings, they concluded that low frequency information within speech plays a very important role in the perceptual segregation of speech from competing background noise.

Recent studies related to hearing devices have identified certain clinical population who show improvement in speech perception when low frequency information is added. Providing low to mid frequency additional amplification to people with severe-to-profound sensori-neural hearing loss has been shown to improve their speech comprehension abilities (Turner & Brus, 2001). Gantz, Turner, Gfeller and Lowder (2007) studied the benefit of hybrid cochlear implant on 21 patients and found a significant improvement in speech understanding in noise over standard cochlear implant participants. They concluded that low frequency hearing is important for speech perception in noise. Advantages of using bimodal implants have been well studied and researches have shown that addition of low frequency information through contralateral hearing aid leads to better perception of speech in quiet, even in noise and improves localization (Mok, Grayden, Dowell & Lawrence, 2006; Ching, Incerti, Hill & Van Wanrooy, 2006; Luntz, Shpak & Weiss, 2005; Hamzavi, Pok, Gstuettner & Baumgartner, 2004; Armstrong, Pegg, James & Blamey, 1997; Shallop, Arndt & Turnacliff, 1992).

Further, it has been reported that frequency importance function for speech perception is language dependant. That is in certain languages, weightage given for low frequencies in understanding speech is more compared to other languages. Avilala, Prabhu and Barman (2010) examined the effect of filtering of monosyllables and words on speech perception on 30 normal adults. They reported that in Kannada, a south Indian language, the low-pass cut off frequency at which 70% speech identification scores were obtained was 1200 Hz for words, which is slightly lower compared to the cut off frequency (1500 Hz), reported for English words (Bornstein, 1994).

Owing to the diversity and the varying underlying pathology of auditory disorders, it is not incorrect to speculate that speech perception of different clinical population is different.

One set of patients may have good speech perception abilities in high frequencies while other in lower frequencies. Few factors underlying their varying capacities may be degree of hearing loss, configuration of hearing loss or underlying pathology itself. For instance, Prabhu, Avilala and Barman (2011) studied the speech perception abilities of individuals with auditory dysynchrony and found that speech understanding scores was poorer for low pass filtered words in comparison to unfiltered words. They also recommended the use of low pass filtered words as a tool for accurate assessment of speech perception difficulties in individuals with auditory dysynchrony.

Speech is highly redundant owing to the simultaneous transmission of the information in several ways. Thus, a hearing loss involving only a part of frequency range may go undetected if the available standard speech tests are used. Conventionally, Phonetically Balanced (PB) word lists are used to assess the speech understanding abilities. PB word lists available in Hindi and Kannada were developed by De (1973) and Yathiraj and Vijayalakshmi (2005) respectively. Usefulness of these tests has been validated considering individuals with flat configuration of hearing loss. Thus, use of these tests for determining the communication problems may not be accurate if an individual has more loss in low or high frequencies.

Several test materials have been developed in Indian languages that can cater to the individuals with high frequency hearing loss. Few of those include, High frequency word list in Hindi (Ramachandra, 2001), Kannada (Kavitha, 2002), English (Sudipta, 2006), Tamil (Sinthiya, 2009) and Telugu (Ratnakar, 2010). These tests are now used to assess the individual's perception of high frequency speech information. However, till date not many attempts have been made to develop speech materials to assess patient's ability to assess low frequency information. Also, the selection of appropriate hearing devices/management option for individuals with low frequency hearing loss calls for the use of a test that exhibits sensitivity to their problems.

All these factors that is a. Importance of low frequency information in speech perception, b. Existence of considerable number of individuals with low frequency hearing loss, c. Existence of hearing disorders with varying speech perception abilities for low and high frequency speech information and d. Unavailability of proper tools to assessing speech perception abilities in low frequency information either for diagnostic purposes or for assessing candidacy for various rehabilitation measures were the reasons behind why this study was taken up.

Objectives of the study:

1. To develop low frequency word lists in Hindi and Kannada languages to obtain speech identification scores.
2. To standardize the test materials by establishing normative data by considering native normal listening individuals in Hindi and Kannada.
3. To check the equity of the lists developed both in Hindi and Kannada.
4. To determine the usefulness of the material by administering it on the individuals with low frequency cochlear hearing loss or by simulating rising type of hearing loss.

METHOD

The present study aimed to develop low frequency word lists in Hindi (Indo-Aryan language of India) and Kannada (Dravidian language of south India). For this purpose the study was conducted in following three phases: Development of the low frequency word lists, standardization of the test materials developed and determining the usefulness of the test materials.

Phase I. Development of the low frequency word lists

This phase was subdivided into 4 stages. They were viz. Collection of words and obtaining familiarity ratings, recording and selection of the best recorded words, separating words with dominant low frequency energy and finally generating word lists with equal difficulty levels.

Collection of words and obtaining familiarity ratings

One of the important factors to be considered while developing speech material is word commonness or familiarity (Hirsh, 1952). Considering this in mind, total of 2012 and 2090 bisyllabic words were collected from common sources in Hindi and Kannada respectively. The words were collected irrespective of the energy concentration of phonemes across frequencies. Those words were verified for the presence of any script errors and correct categorization as bisyllabic words by respective native linguists. It was then evaluated by a respective native linguist for script errors and correct categorization as bisyllabic words. Further, a 5 point familiarity rating scale ('unfamiliar', 'less familiar', 'familiar', 'more familiar' & 'most familiar') was used to get familiarity ratings for all words by 10 native adult speakers. 'Less familiar' and 'unfamiliar' words were eliminated and the words left over were 1319 and 1285 in Hindi and Kannada respectively.

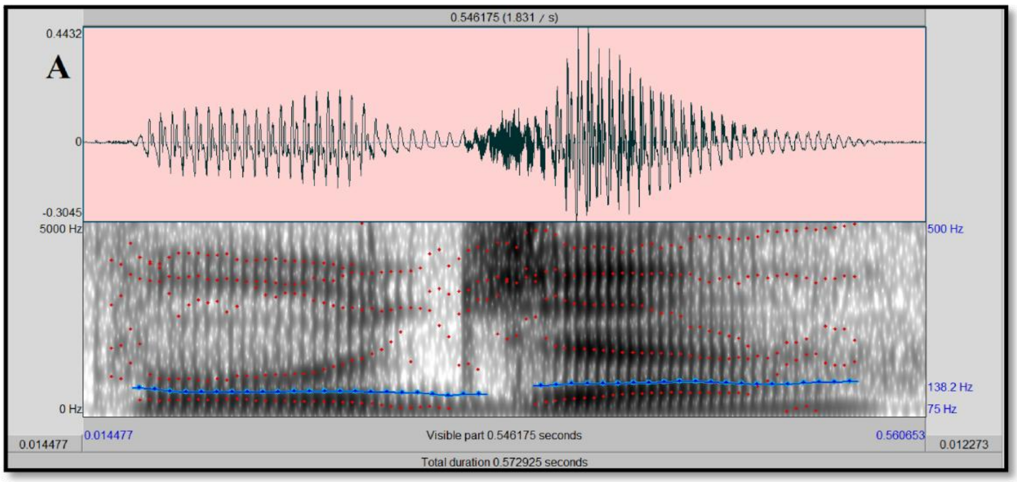
Familiarity ratings by all the participants were complied and analyzed. The words rated familiar, more familiar or most familiar by 70% of the participants were considered and rests of the words were excluded from the list.

Recording of words and selection of best recorded words

Selected words were recorded in a sound treated room using Computerized Speech Lab - Model 4500, an input/output recording device for the personal computer by KayPENTAX. Recording was done using 16 bit analogue to digital converter and at a sampling rate of 44.1 kHz. Unidirectional dynamic microphone (Shure SM-48) was used for the recording and was kept at a distance of around 6 cm from speaker's mouth.

Initially, 15 words from selected words were recorded considering 3 native adult male fluent speakers of Hindi and Kannada. As, mean fundamental frequency of male speakers is lower compared to female speakers, male voice was preferred for recording the words. These voice recordings were then given to 4 adult native speech and hearing professionals with normal hearing for judging the most appropriate voice to record the entire list. Judges were asked to rate the voice based on parameters like voice quality, clarity and naturalness on a 3- point rating scale (poor, fair and good). One of the recorded voices, which received the highest score ratings was selected and that person's voice was used for recording all the words selected.

Selected familiar words were subsequently recorded with selected speaker's voice and the entire list was recorded in 5 different sessions. Each word was recorded five times in clear and monotonous voice. Out of five recordings, first and the last recordings were removed and only the middle three were subjected to subjective and objective analysis to select the best recorded words. Pratt software (version 5.3.03) was used for objective analysis. Firstly, words were subjectively analyzed and rated by an experienced audiologist for the clarity of the utterance, presence of any intonation patterns and background audible noise. Out of three repetitions of each word, the best rated recordings, which were free of background noise, clear and monotonous were considered. Further, among those recordings, one with visible pitch and formants observed using Pratt software was finally selected during objective analysis. Those words for which all five repetitions got eliminated during subjective and objective analysis as they did not satisfy one or many of the criteria during analysis were re-recorded. The entire analysis procedure was repeated for the re-recorded words. The recordings of the selected words were subjected to intensity normalization using MATLAB software R2009b. Figure.1 depicts the spectrogram of a sample word in Hindi and Kannada.



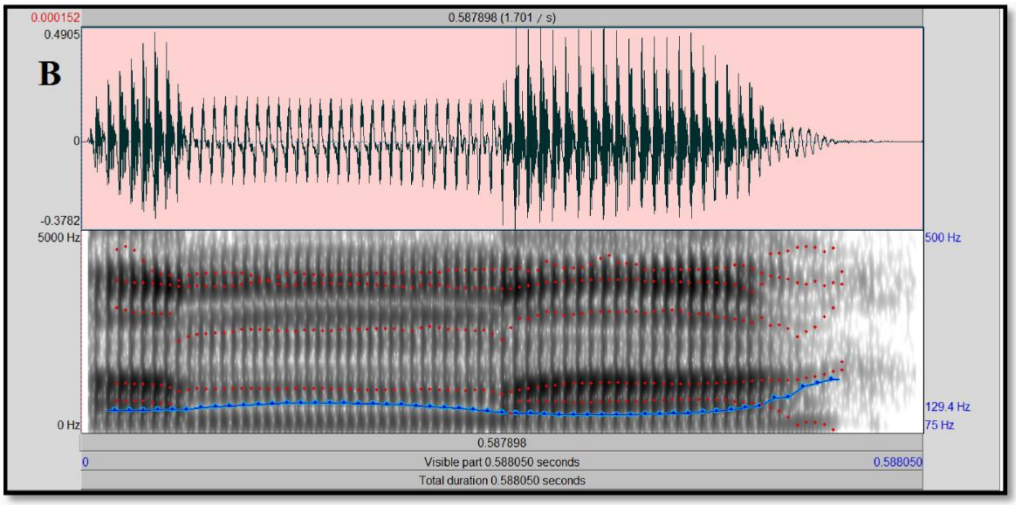


Figure 1: Spectrogram of words A. Hindi (/pooja/) and B. Kannada (/amma/) showing formants (red dotted line) and pitch (blue solid line).

Separating words with dominant low frequency energy

In order to separate words with predominantly low frequency energy from rest of the words, using MATLAB software, FFT was performed to determine each word's energy below and above 1.5 kHz. According to Maltby and Knight (2000) and Warner Brown and McCartney (1984), the important speech frequencies fall between 250 Hz to 4000 Hz. 1500Hz cut off frequency was selected as it falls around center of this frequency range. Later, Amplitude ratio; the ratio of energy below and above 1.5 kHz was obtained. Words with higher amplitude ratio are those with more low frequency energy. Thus, to separate those low frequency words, *k*-means clustering was carried out based on the amplitude ratio of the words. Amplitude cut off ratios for Hindi and Kannada were 1.8 and 1.9 respectively. This resulted in separating low frequency clusters of words from others. A total of 850 words in Hindi and 829 words in Kannada were grouped as low frequency words. To verify the difference in energy distribution across frequencies, Long Term Average Speech Spectrum (LTASS) was executed on the clusters of words grouped as low frequency. Figure 2 illustrates LTASS results. The difference in energy between low and high frequency region is clearly evident in the figure.

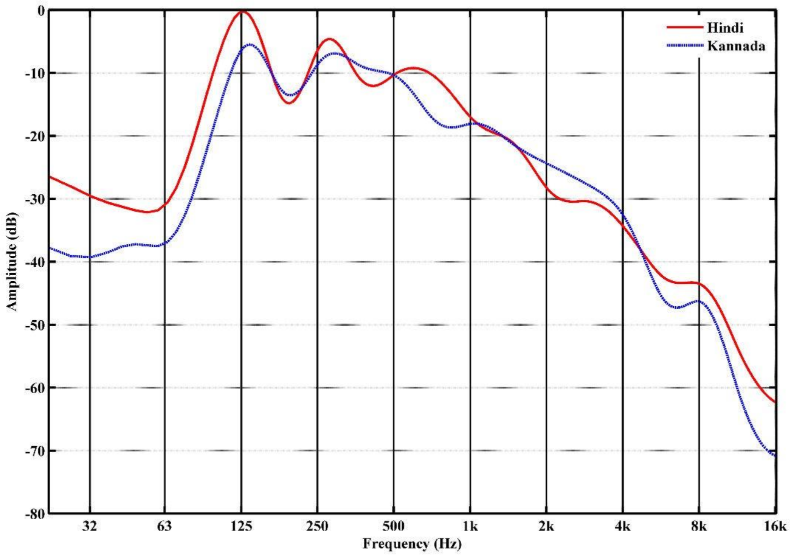


Figure 2: LTASS results for low frequency cluster of words in Hindi and Kannada.

To further ensure that the lists do not include phonemes with high frequency energy, the words were also subjected to phonemic analysis. During which, attempts were made to remove words with consonants considered as high frequency both in Hindi and Kannada. Subsequently, only 721 words in Hindi and 498 words in Kannada were left over.

Generating word lists with equal difficulty

All the low frequency words were then presented to 25 adult native speakers with normal hearing at 5 different sensation levels (ref: PTA). SLs considered were +0, +4, +8, +12 and +16 dB. A calibrated dual channel diagnostic audiometer, MAICO 53, was used for screening participants hearing and also to present the words. The recorded speech material was played using MATLAB software R2009b. The signal was routed through a personal computer to the audiometer and presented through headphones, Sennheiser HD-200.

Further, speech identification (SI) scores were calculated using the following formula:

$$\text{SI Score} = \frac{\text{Totalnumberofcorrectresponses}}{\text{Total number of words presented}} \times 100$$

The SI scores obtained from the participants at each SL were averaged and tabulated. Based on the averaged scores at all SLs, psychometric functions were derived for all the words using MATLAB software R2009b. Mean sensation level where 50% SI scores occurred and mean slope of the psychometric functions were obtained. Words falling within ±1.5 standard deviation from overall mean and slope were accepted. These words were used to make the final word lists of 25 words each. However, lists were not phonemically balanced as high frequency consonants were removed. This consisted of 691 words in Hindi and 247 words in Kannada.

Using this word pool, final Kannada and Hindi word lists were formed. These word lists were generated such that all lists had equal difficulty levels. Equality in difficulty was verified by firstly randomly selecting 25 words from 691 Hindi and 247 Kannada words, generating psychometric function curves for those words. Further, mean level at which 50% SI scores were obtained and slope of the curve was found. If the mean level and slope fell within ± 1.5 standard deviation from overall mean and slope obtained earlier for 721 Hindi and 498 Kannada words, then those 25 words chosen randomly formed one final word list.Same procedure was repeated for forming rest of the lists. Finally, a total of 10 lists in Hindi and 7 lists in Kannada were developed. Hence, these word lists developed can be called as psychometrically equivalent word lists. These lists were not phonetically or phonemically balanced as all the phonemes did not appear in the lists. Appendix I provide the details of word lists developed finally in Hindi and Kannada.

Phase II. Standardization of the test material

The developed test materials in Hindi and Kannada were standardized by obtaining SI scores from 100 adult native Hindi and Kannada speakers each. All the participants were in the age range of 18 to 55 years. It was ensured that all the participants had hearing thresholds within 15dB HL at octave frequencies between 250Hz and 8000Hz and SI scores above 90% (Yathiraj & Vijaylakshmi, 2005; De, 1973). The participants had type 'A' tympanogram with acoustic reflexes present bilaterally. Also, all the participants had bilateral TEOAEs present and more than 60% score on speech perception in noise test at 0 dB signal to noise ratio.

A calibrated dual channel diagnostic audiometer, MAICO 53, was used to carry out pure tone and speech audiometry. The recorded speech material was played using MATLAB software R2009b. The signal was routed through a personal computer to the audiometer and presented through headphones, Sennheiser HD-200.

The word lists developed in Phase I were used to obtain SI scores at 0, +10, +20, and +30 dB SL (ref: PTA) in 10 participants at each level. To standardize the material, all the word lists were presented at 40 dB SL (ref: PTA) to 100 participants in Hindi and Kannada. To avoid ear effect, for 50 individuals words were presented to the right ear and for remaining 50 participants left ear was considered for testing. An open set response in the form of verbal repetitions was obtained from all the participants. The SI scores at each SL were averaged, tabulated and analyzed using the software Statistical Package for Social Sciences (SPSS), Version 20.

Phase III: Determining the usefulness of the test material

After standardization of the test, the standardized Kannada lists were presented to 10 adult native language speakers with cochlear hearing loss (rising pattern) to obtain the SI scores. The categorization of rising type was made based on the classification given by Lloyd and Kaplan (1978). Due to non-availability of Hindi speaking patients with low frequency cochlear hearing loss at the centre where data was collected, rising cochlear hearing loss was simulated on normal hearing individuals using MATLAB R2009b and NIOSH Hearing Loss Simulator software (version 3.0.12151).

Table 1: Demographic data of Hindi speaking individuals with simulated rising cochlear HL

	Age (years)/ Gender	**Ear**	**Pure Tone Average (dB HL)**	**SIS (%)**
1	18/M	Rt	10	100
2	20/F	Rt	11.25	100

3	18/F	Lt	10	96
4	19/M	Rt	12.5	100
5	20/F	Lt	7.5	100
6	22/F	Lt	6.25	96
7	25/M	Rt	15	100
8	19/F	Lt	5	100
9	28F	Rt	7.5	100
10	21/M	Lt	11.25	100

Note: SIS- Speech Identification Scores, Lt-Left, Rt-Right

Table 2: Demographic data of Kannada speaking patients with low frequency cochlear HL

	Age (years)/ Gender	**Ear**	**Pure Tone Average (dB HL)**	**SIS (%)**
1	60/M	Rt	35	100
2	21/M	Lt	35	92
3	27/M	Lt	33.75	100
4	60/F	Rt	36.25	92
5	30/M	Rt	52.5	80
6	18/M	Lt	40	100
7	35/M	Lt	45	88
8	48/M	Lt	50	100
9	20/M	Lt	35	92
10	42/M	Rt	28.75	100

Note: SIS- Speech Identification Scores, Lt-Left, Rt-Right

Using NOISH software, moderate rising low frequency hearing loss was simulated for Hindi Speakers by manually setting the hearing thresholds for octave frequencies between 125 Hz to 8000 Hz. Using MATLAB software reduced frequency selectivity was added to the hearing loss simulation. The procedure followed to generate spectrally smeared output was described by Moore and Glassberg (1993). Schematic diagram of the sequence of operations used to perform spectral smearing is shown in the Figure 3. That involved firstly calculating short term spectrum using a Hamming window and an FFT. Followed by spectrum was smeared and then it was transformed back to time domain using inverse FFT. At last, waveforms obtained from overlapping analysis frames were added to produce final output.

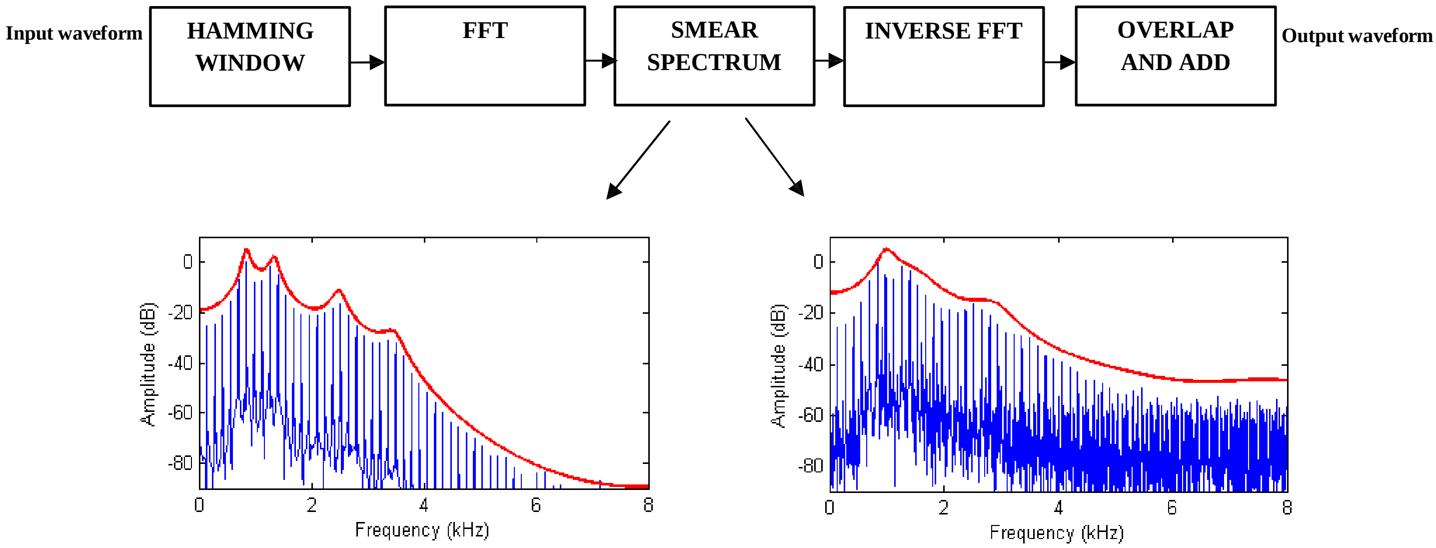


Figure 3: Schematic diagram of the sequence of operations used to perform spectra smearing

Age range of the Kannada and Hindi participants was between 18-60 years and 18-28 years respectively. The degree of hearing loss of the Kannada participants varied from mild to moderate hearing loss. Hearing thresholds were obtained for the octave frequencies between 250 to 8000Hz. Pure Tone Average (PTA), that is average of thresholds at 0.5k, 1k, 2k and 4k was obtained. Also, SI scores were obtained using phonetically balanced word lists available in Hindi and Kannada (De, 1973; Yathiraj & Vijayalakshmi, 2005). Tables 1 and 2 provide demographic data of the participants.

The standardized low frequency word lists (10 lists in Hindi and 7 in Kannada) were presented to 100 participants at 40 dB SL (ref. PTA). The presentations were randomized, both in terms of order of words and lists. SI Scores were obtained using the formula given in Phase I. Further, scores were averaged and tabulated for all the lists separately. To explore the difference in performance between normals and individuals with rising cochlear hearing loss data was analyzed using the SPSS software (Version 20).

RESULTS AND DISCUSSION

This study was carried out in three phases 1. Development of the low frequency word lists. 2. Standardization of the word lists developed and 3. Validating the usefulness of the test material developed. Results of all the phases are provided separately in detail below.

1. Development of the low frequency word lists

This phase involved 4 stages viz. a. Collection of words and obtaining familiarity ratings, b. Recording and selection of the best recorded words, c. Separating words with dominant low frequency energy and d. Generating word lists with equal difficulty levels. The procedure for collection of words and recording, selection of best recorded words and separating words with dominant low frequency energy are explained in Method section.

Generating word lists with equal difficulty levels

One of the important considerations during development of any speech material for testing is that the alternative forms of testing should be equivalent that is they should produce comparable results (Gelfand, 2009). Conventional way followed to obtain equivalency between word lists developed was phonemic or phonetic balancing. However, recent researches in the field of development of speech identification materials have shown that impact of phonetic or phonemic balancing is questionable (Martin, Champlin & Perez, 2000). Thus, to develop word lists which produce equivalent results, following procedure considering psychometric function curves was carried out.

Using low frequency word cluster, the results of speech identification (SI) scores obtained from 25 adult native speakers with normal hearing at sensation levels of +0, +4, +8, +12 and +16 dB was analyzed. The SI scores were calculated, averaged and tabulated. Based on the average scores at all SL's, the psychometric function curves were obtained for all the words. Based on these psychometric function curves, mean level at which 50% SI score was obtained and the slope of function was derived. The words falling within ±1.5 standard deviation from the mean and the slope for the psychometric functions were separated. This consisted of 691 words in Hindi and 247 words in Kannada. Based on the procedure explained in method section, a total of 10 lists in Hindi and 7 lists in Kannada were developed.

2. Standardization of the test material

The mean and SD of SI scores obtained at 40 dB SL in 100 individuals with normal hearing for 7 lists in Kannada is shown in figure 4 and 10 lists in Hindi are shown in figure 5. Repeated measures ANOVA was done to compare differences across lists for Kannada and Hindi separately. The results showed that there was no significant difference across lists for both Kannada [$F(6,594) = 37.86$, $p>0.05$] and Hindi [$F(9,891) = 34.93$, $p>0.05$] at 40 dB SL.

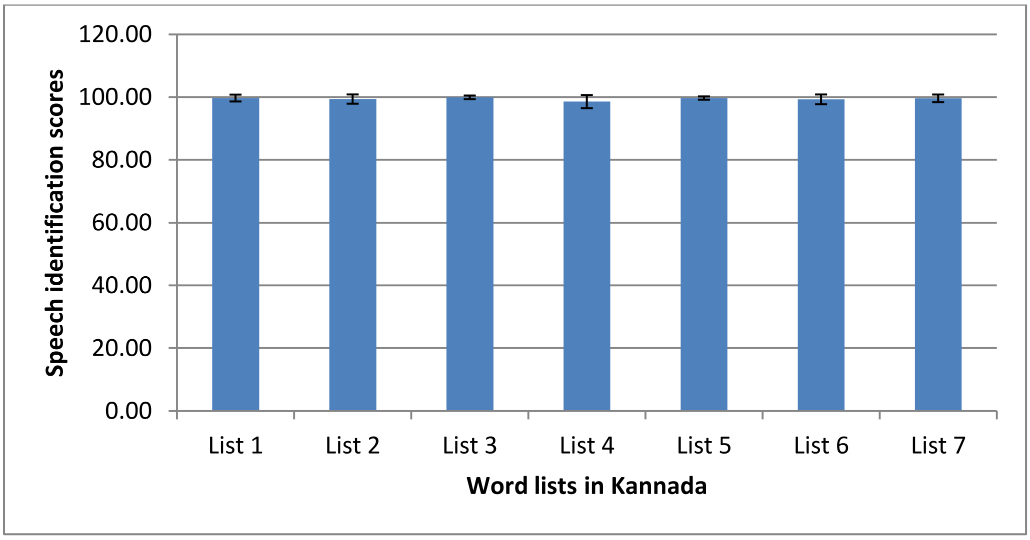


Figure 4: Mean and SD of SI scores obtained for word lists in Kannada at 40 dB SL.

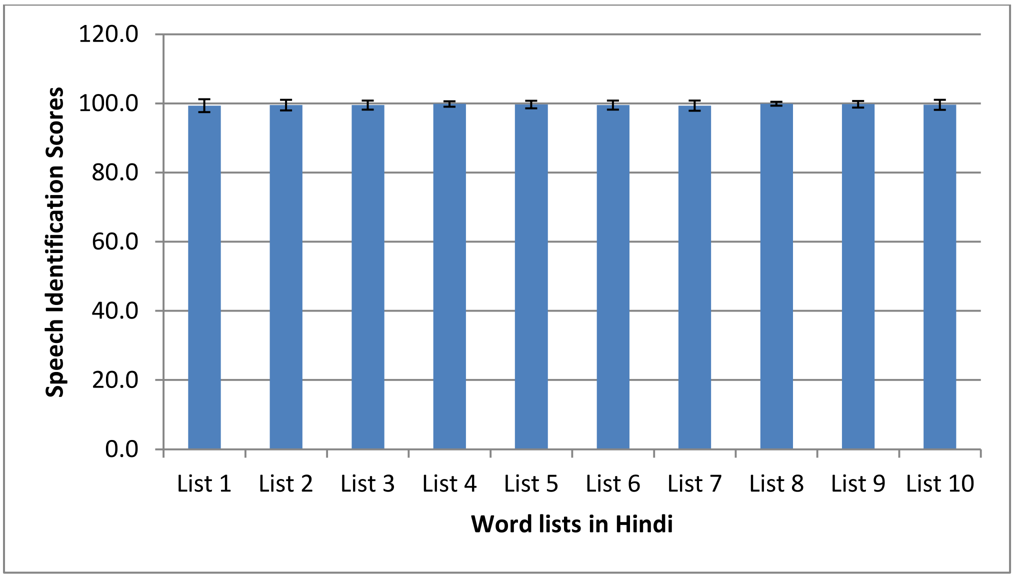


Figure 5: Mean and SD of SI scores obtained for word lists in Hindi at 40 dB SL.

The SI scores were also obtained from 40 participants for the final lists across 0, +10, +20 and +30dB SL (ref. PTA). Figures 6 and 7 depict mean SI scores and standard deviations for all the lists across SLs in Hindi and Kannada respectively.

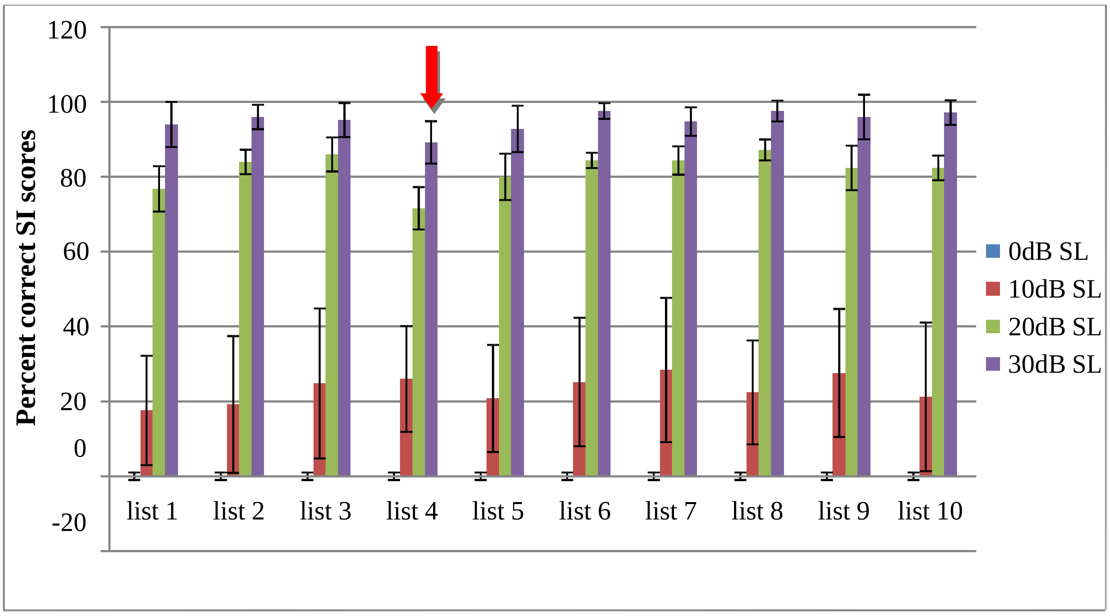


Figure 6: Mean SI scores and standard deviations of all the lists across SLs in Hindi. Red arrow in the graph highlights the finding that, SI scores of the list 4 were lower compared to other lists at 20 and 30dB SLs.

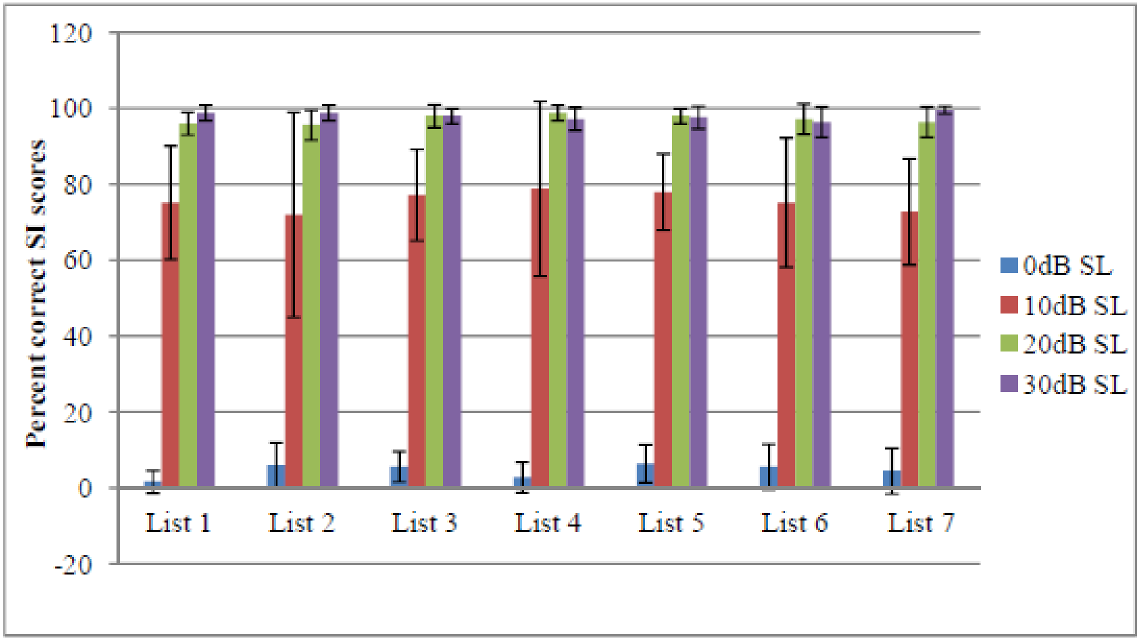


Figure 7: Mean SI scores and standard deviations of all the lists across SLs in Kannada.

To inspect equality of lists, psychometric function curves were obtained for SI scores across SLs in Hindi and Kannada (Figure 8). It was found that except for one list in Hindi all the other lists were following the same trend in difficulty across SLs. In psychometric function curves of Hindi, it can be noted that 100% scores was reached almost at 30dB SL. This is in congruence with earlier reports where maximum speech identification scores were obtained in normal at 30 to 40 dB SL (Gold, Lubinsky & Shahar 1981). In Kannada, it can be observed that all lists are following almost similar trend but unlike Hindi maximum scores were obtained at lower SLs itself (around 20dB SL).

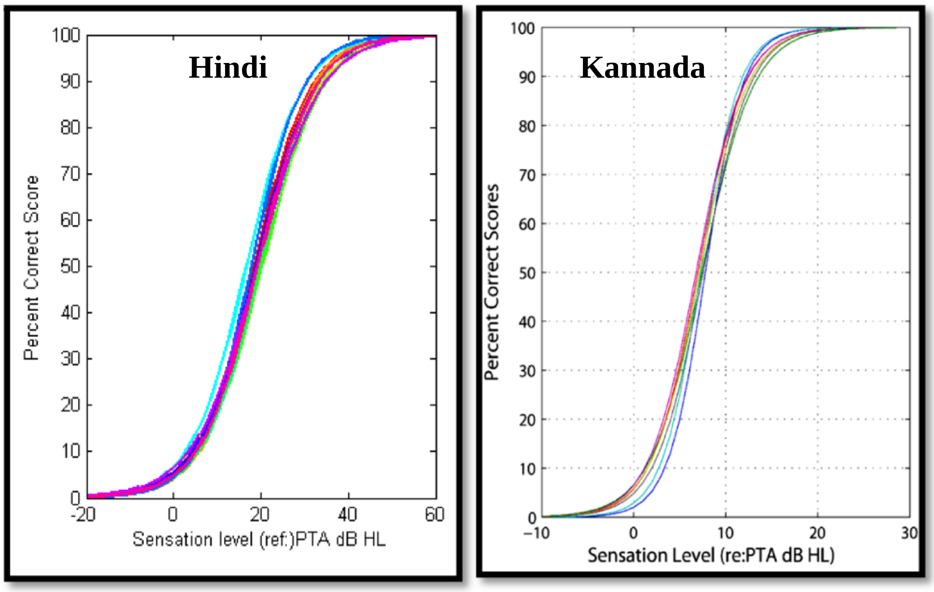


Figure 8: Psychometric function curves across SLs for wordlists in Hindi and Kannada.

In the Figures 4 and 5, it can be observed that, in general SI scores are increasing with the increase in the intensity for all the lists. To study the statistical difference, data was analyzed using SPSS software. Shapiro Wilks test of normality was done and it showed normal distribution till 20 dB SL except at 30 dB SL. As the data obtained at 3 SLs were normally distributed the mean scores obtained for the word lists were compared across SLs using repeated measures ANOVA. It was noticed that, for Hindi word lists, there was a significant difference between SLs [$F(3, 27) = 302.42$, $p<0.001$], and even across lists [$F(9, 81) = 4.474$, $p<0.001$]. Interaction between lists and SLs were also found to be significant [$F(27, 243) = 2.483$, $p<0.001$].

Post hoc analysis was done using Sidak test and the results revealed that, scores across SL's were significantly different from one another, with scores increasing from 0dB SL to 30dB SL ($p<0.001$). At 0dB SL and 10dB SL, there was no significant difference between SI scores across lists, while at 20dB SL, scores obtained in list 4 was found to be significantly lower from list 2 and 8. Even at 30dB SL, SI scores of word list 4 was significantly lower from lists 2, 6, 8 and 10 ($p<0.05$). Probable reason for such a finding was poorer SI scores of list 4 at

20 and 30dB SL (Figure 3) compared to other lists may be due to its higher difficulty. It is possible that the list had more number of words that were less familiar as 70% word familiarity criteria was used. Hence, it is suggested not to use this list for clinical purposes. Instead the words in this list may be used as practice material.

In Kannada, repeated measures ANOVA also showed a significant difference between SLs [F (3, 27) = 325.15, p<0.001]. Further, ignoring SLs, across list there was no significant difference in the scores [F (6, 54) = 27.84, p>0.05] and no significant interaction between SLs and lists [F (18, 162) = 32.267, p>0.05]. Post hoc analysis was done using Sidak test. Results revealed that, scores of 0dB and 10dB SL were significantly different from the scores of 20 and 30dB SL. Also, scores of 0dB SL was significantly different from that of 10dB SL (p<0.001). While scores of 20dB and 30dB SL were not significantly different from each other. In other words, SIS scores increased from 0 to 20dB SL and stabilized thereafter.

Table 3: Post hoc analysis of variation across lists in Hindi and across SLs for Kannada word list

Hindi Word lists								Kannada Word list	
0 dB SL		10 dB SL		20 dB SL		30 dB SL			
List 1	List 2	List 1	List 2	List 1	List 2	List 1	List 2	0 dB SL	10 dB SL
	List 3		List 3		List 3		List 3		20 dB SL*
	List 4		List 4		List 4		List 4		30 dB SL*
	List 5		List 5		List 5		List 5	10 dB SL	20 dB SL*
	List 6		List 6		List 6		List 6		30 dB SL*
	List 7		List 7		List 7		List 7	20 dB SL	30 dB SL
	List 8		List 8		List 8		List 8		
	List 9		List 9		List 9		List 9		
	List 10		List 10		List 10		List 10	*Significant difference with p<0.05.	
List 2	List 3	List 2	List 3	List 2	List 3	List 2	List 3		
	List 4		List 4		List 4*		List 4*		
	List 5		List 5		List 5		List 5		
	List 6		List 6		List 6		List 6		
	List 7		List 7		List 7		List 7		
	List 8		List 8		List 8		List 8		
	List 9		List 9		List 9		List 9		
	List 10		List 10		List 10		List 10		
List 3	List 4	List 3	List 4	List 3	List 4	List 3	List 4		

	List 5		List 5		List 5		List 5
	List 6		List 6		List 6		List 6
	List 7		List 7		List 7		List 7
	List 8		List 8		List 8		List 8
	List 9		List 9		List 9		List 9
	List 10		List 10		List 10		List 10
List 4	List 5	List 4	List 5	List 4	List 5	List 4	List 5
	List 6		List 6		List 6		List 6*
	List 7		List 7		List 7		List 7
	List 8		List 8		List 8*		List 8*
	List 9		List 9		List 9		List 9
	List 10		List 10		List 10		List 10*
List 5	List 6	List 5	List 6	List 5	List 6	List 5	List 6
	List 7		List 7		List 7		List 7
	List 8		List 8		List 8		List 8
	List 9		List 9		List 9		List 9
	List 10		List 10		List 10		List 10
List 6	List 7	List 6	List 7	List 6	List 7	List 6	List 7
	List 8		List 8		List 8		List 8
	List 9		List 9		List 9		List 9
	List 10		List 10		List 10		List 10
List 7	List 8	List 7	List 8	List 7	List 8	List 7	List 8
	List 9		List 9		List 9		List 9
	List 10		List 10		List 10		List 10
List 8	List 9	List 8	List 9	List 8	List 9	List 8	List 9
	List 10		List 10		List 10		List 10
List 9	List 10	List 9	List 10	List 9	List 10	List 9	List 10

2. Determining the usefulness of the test material

To determine the usefulness of the test material, the lists in Kannada were administered on 10 native individuals with rising cochlear hearing loss and 10 Hindi speaking individuals with simulated condition resembling rising cochlear hearing loss. Figures 9 and 10 show the mean and standard deviation of SI scores between normals and simulated low frequency cochlear hearing loss (SLCHL) in Hindi and between normals individuals with low frequency cochlear hearing loss (LCHL) in Kannada respectively. It is clear from the figures that across lists, normals have outperformed LCHL and SLCHL individuals in Kannada and Hindi respectively.

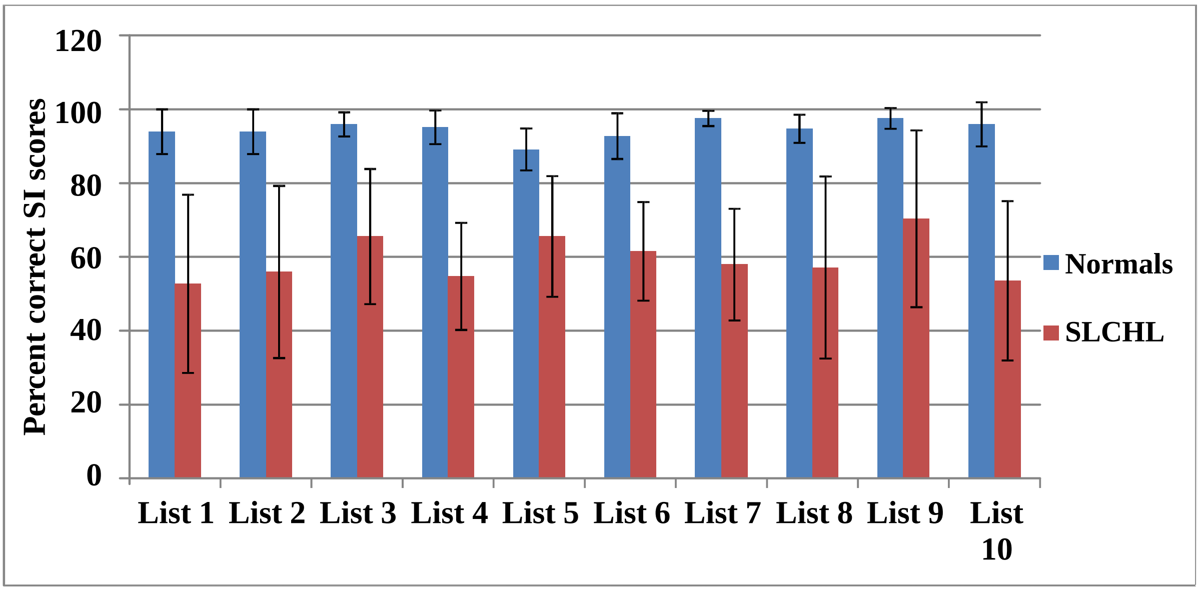


Figure 9: Mean and SD of SI scores between normals and SLCHL group in Hindi.

Shapiro Wilks test of normality was done and it showed normal distribution for normal hearing, LCHL and SLCHL for all the word lists in Kannada and Hindi. Levene's test of homogeneity of variances was done and the variances were homogenous across the groups. Mixed ANOVA was carried out to find if there was any significant difference in SI scores between groups across word lists in Hindi and Kannada separately. In Hindi, main effects showed a significant difference between lists [F (9, 162) = 2.064, $p<0.05$] and between normal and hearing impaired groups [F (1, 18) = 50.076, $p<0.001$]. Also, significant interaction was seen between lists and groups [F (9, 162) = 2.188, $p<0.05$]. In Kannada, main effects showed significant difference between groups [F (1, 18) = 7.959, $p<0.05$] but not across lists [F (6,108) = 0.692, $p>0.05$]. Interaction effects were significant [F (6,108) = 2.405, $p<0.05$]. Post hoc analysis using Sidak test showed that only list 4 was significantly different from list 3 and 6 in Hindi.

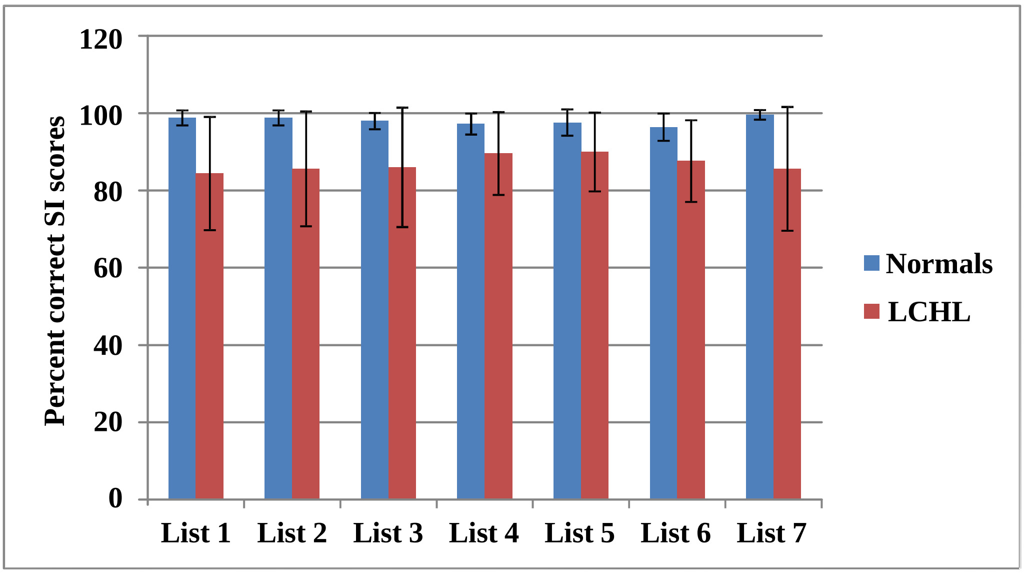


Figure 10: Mean and SD of SI scores between normals and LCHL group in Kannada.

Both in Hindi and Kannada, word lists were found to be difficult for individuals with rising cochlear hearing loss, thus validating the usefulness of the lists developed. However, there was more reduction in mean scores in Hindi SLCHL group compared to Kannada (Figure 9), this could be due to variation in the group recruited to validate the lists. That is in Kannada, individuals with rising cochlear hearing loss were considered while in Hindi lists were administered on normal hearing individuals by simulating rising cochlear hearing loss. Hence, the difference could be due to various possible factors. One of those could be internal physiological differences. Another reason could be difference in degree of hearing loss. In Hindi group, moderate degree of hearing loss was simulated for all the participants, while in Kannada, patient group had hearing loss ranging between mild to moderate. However, 7 out of 10 Kannada participants had mild hearing loss and thus, lesser degree of hearing loss could have lead to better scores in Kannada speaking group compared to Hindi participants.

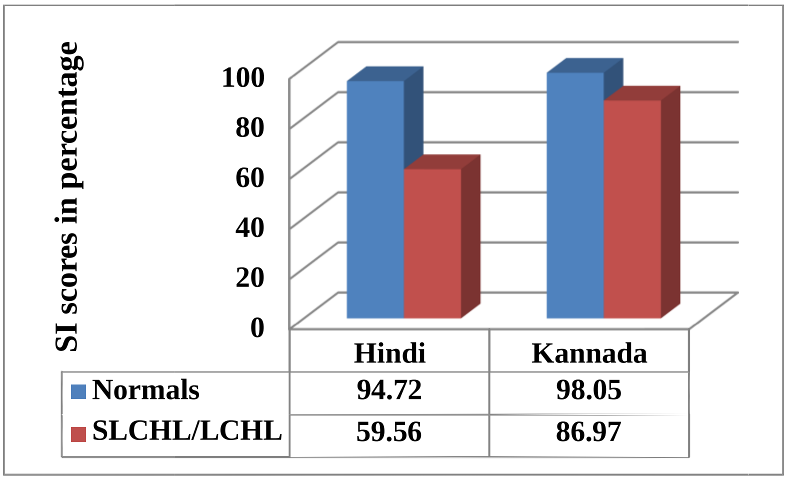


Figure 11: Overall mean SI scores for normal and SLCHL/LCFL group in Hindi and Kannada.

CONCLUSION

This study resulted in development and standardization of low frequency word lists in Hindi and Kannada languages. In total, 10 word lists in Hindi and 7 word lists in Kannada were developed. However, one list (list 4) in Hindi was found to vary from other lists and thus it is suggested not to use this list for clinical purposes but can be used as practice list. The present study also validated the list on low frequency cochlear hearing loss individuals and recommends the use of these standardized word lists on clinical population to tap their difficulty in understanding low frequency information. Further, studies can be taken up to validate the use of this speech material on other clinical population also.

REFERENCES

Armstrong, M., Pegg, P., James, C., & Blamey, P. (1997). Speech perception in noise with cochlear implant and hearing aid. *American Journal of Otology, 18,* S140-S141.

Baer, T., & Moore, B. C. J. (1993). Effects of spectral smearing on the intelligibility of sentences in noise, *Journal of Acoustical Society of America, 94(3),* 1229-1241.

Ching, T., Incerti, P., Hill, M., & Van Wanrooy, E. (2006b). An overview of binaural advantages for children and adults who use binaural/bimodal hearing devices. *Audiology Neurotology, 11* (suppl 1), 6-11.

Dobie, A. R., & Van Hemel, S. B. (2004). *Hearing loss: Determining Eligibility for social security benefits.* Washington, D. C, National Academies Press.

Eric, T., Benedicte, P., Jean-Francois, V., Samia, L., & Lionel, C. (2008). Vibrant Soundbridge Versus Conventional Hearing Aid in Sensorineural High-Frequency Hearing Loss, A Prospective Study, *Otology & Neurotology, 29(5)* 684-687.

Gardner, H. J. (1971). Application of a High-Frequency Consonant Discrimination Word List in Hearing-Aid Evaluation. *Journal of Speech and Hearing Disorders, 36,* 354-355.

Gates, G. A., Couropmitree, N. N., & Myers, R. H. (1999). Genetic associations in age-related hearing thresholds. *Archives of Otolaryngolgy - Head & Neck Surgery. 125,* 654–659.

Gelfand, S. (2009). *Essentials of Audiology*. (3rd Edition). New York, NY: Thieme.

Hamzavi, J., Pok, S., Gstoettner, W., & Baumgartner, W., (2004). Speech perception with a cochlear implant used in conjunction with a hearing aid in the opposite ear. *International Journal of Audiology, 43,* 61-66.

Hannley, M, T. (1993). Audiological characteristics of the patients with otosclerosis. *Otolaryngologic Clinics of North America, 26,* 373-87.

Harner, S. G., Fabry, D. A., & Beatty, C. W. (2000). Audiometric findings in patients with acoustic neuroma. *American Journal of Otolaryngology, 21*(3), 405-411.

Hull, F. M., Mielke, P. W., Jr., Timmons, R. J., & Willeford, J. A. (1971). The national speech and hearing survey: Preliminary results. *Asha, 13*(9), 501–509.

Jin, S. H., & Nelson, P. B. (2010). Interrupted speech perception: the effects of hearing sensitivity and frequency resolution. *Journal of Acoustic Society of America, 128*(2), 881–889.

Kavitha, E. M., & Yathiraj, A. (2002). High Frequency-Kannada Speech Identification Test (HF-KST), *Unpublished Master's dissertation*, Submitted to University of Mysore, Mysore.

Knox, G.W., & McPherson, A. (1997). Meniere's disease: differential diagnosis and treatment. *American Family Physician*, *55,* 1185–1190.

Lloyd, L.L., & Kaplan, H. (1978). *Audiometric interpretation.* Baltimore, MD: University Park Press.

Luntz, M., Shpak, T., & Weiss, H. (2005). Binaural-bimodal hearing: Concomittant use of a unilateral cochlear implant and a contralateral hearing aid. *Acta Oto-Laryngolica, 125,* 863-869.

Maltby, M.A. & Knight, P. (2000). Audiology: an introduction for teachers and other professionals. David Fulton Publishers: New York, Pp 14.

Margolis R. H., Saly G., (2008). "Asymmetric hearing loss: definition, validation, and prevalence," *Otology & neurotology, 29*(4), 422–431.

Mendel, L.L. & Danhauer, J.L. (Eds.) (1997). *Audiologic Evaluation and Management and Speech Perception Assessment.* San Diego, CA: Singular Publishing Company, Inc.

Mok, M., Grayden, D., Dowell, R., & Lawrence, D. (2006). Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. *Journal of Speech, Language, and Hearing Research, 49*, 338-351.

Noble, W., Sinclair, S., & Byrne, D. (1998). Improvement in aided sound localization with open earmolds: observations in people with high-frequency hearing loss. *Journal of American Academy of Audiology, 9* (1), 25-34.

Pascoe, D, P. (1975). Frequency response of hearing aids and their effects on the speech perception of hearing-impaired subjects. *Annals of Otology, Rhinology & Laryngology*, *84* (Suppl. 23), 1-40.

Prabhu, P., Avilala, A. K. Y., & Barman, A. (2011). Effect of Filtering on Perception of Monosyllables (CV) and Words. *Asia Pacific Journal of Speech, Language and Hearing, 14*(2), 111-118.

Rabinowitz, P. M., Slade, M. D., Galusha, D., Dixon-Ernst, C., & Cullen, M. R. (2006b). Trends in the prevalence of hearing loss among young adults entering an industrial workforce 1985 to 2004, *Ear and Hearing, 27*(4), 369–375.

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

Rathnakar, Y. V., & Mamatha, N. M. (2010). High Frequency Speech Identification Test in Telugu, *Unpublished Master's dissertation*, Submitted to University of Mysore, Mysore.

Warner, J., Brown, B.B. & McCartney, E. (1984). Speech Therapy: A Clinical companion. Manchester University Press: Manchester Pp 52.

Shallop, J. K., & Arndt, P. L., & Turnacliff, K. A. (1992). Expanded indications for cochlear implantation: Perceptual results in seven adults with residual hearing. *Journal of Spoken Language, Pathology, Audiology*. *16*, 141–148.

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

Sinthiya, K., & Sandeep, M. (2009). High Frequency Speech Identification Test in Tamil, *Unpublished Master's dissertation*, Submitted to University of Mysore, Mysore.

Stelmachowicz, P, G., Pittman, A, L., Brenda, M. H., Lewis, E, D., & Moeller, M, P. (2004). The Importance of High-Frequency Audibility in the Speech and Language Development of Children with Hearing Loss. *Archives of Otolaryngolgy - Head & Neck Surgery*. *130(5)*, 556-562.

Sudipta, K. B., & Yathiraj, A. (2006). High Frequency-English Speech Identification Test (HF-ESIT), *Unpublished Master's dissertation*, Submitted to University of Mysore, Mysore.

Tarter, S. K., & Robins, T. G. (1990). Chronic noise exposure. High frequency hearing loss and hypertension among automotive assembly workers. *Journal of Occupational Medicine, 32,* 685–689.

Turner, C.W., & Brus, S.L. (2001). Providing low- and mid-frequency speech information listeners with sensorineural hearing loss. *Journal of the Acoustical Society of America, 109*(6), 2999-3006.

Wang, J., Duan, J., Li, Q., Huang, X., Chen, H., Jin, J., Gong, S & Kong, W. (2002). Audiological characteristics of auditory neuropathy, *Zhonghua Er Bi Yan Hou Ke Za Zhi, 37*(4), 252-255.

KANNADA FINAL WORD LISTS

Sl No.	List 1	List 2	List 3	List 4	List 5	List 6	List 7
1a	ಹೂವು	ಮಾವು	¨sÀæªÉÄ	ನೋಡು	ತಾಳ್ಮೆ	ಮಂತ್ರಿ	ಮಡಿ
2b	ಮೋಡ	ನೌಕೆ	ಮಗು	ಪ್ರಾಪ್ತಿ	ಗೌರಿ	ಅಲ್ಪ	ತಲೆ
3c	ಮೌಲ್ಯ	ದಳ	ಧನ	ಕ್ರೌರ್ಯ	ಕೊಡು	ದಿಕ್ಕು	ಕೋಣ
4d	ಹಲ್ಲು	ಮೂಳೆ	ಆಟ	ಓದು	ಗೋಡೆ	ತೆಪ್ಪ	ಆಳ
5e	ನೆರೆ	ಯೋಧ	ವಾರ	ವಾರ್ತೆ	ರಥ	ಅವು	ಕೋಪ
6f	ಲೋಕ	ದಪ್ಪ	ಒಬ್ಬ	ಕಣ್ಣು	ರೊಕ್ಕ	ನಿದ್ದೆ	ಮೊಲ
7g	ಪಾಠ	ಪುಣ್ಯ	ರೂಪ	ಕ್ರೋಧ	ಅರ್ಧ	ಒಟ್ಟು	ಒಪ್ಪು
8h	ವಧು	ಹಾರ	ಹುಲ್ಲು	ಓಟ	ರಕ್ತ	ಕಾಡು	ಮೇಳ
9i	ಮೊರ	ಆಪ್ತ	ರುಧ್ರ	ಮೊಟ್ಟೆ	ಕಾರ್ಯ	ಲೋಪ	ಕ್ರಮ
10j	ಹುಟ್ಟು	ಪಾತ್ರೆ	ನರಿ	ಕೋಣೆ	ಯಾವ	ವೇದ	ಗೆರೆ
11k	ಆರು	ಕೋಳಿ	ದಿವ್ಯ	ಯಾರು	ಪುತ್ರಿ	ಆಳು	ಬೆಟ್ಟ
12l	ಬೇರೆ	ಬೆನ್ನು	ನೀನು	ಟೋಪಿ	ಹೊಲ	ವಾಕ್ಯ	ದಿನ
13m	ದೂರ	ಹಾರು	ನಿತ್ಯ	ಧರ್ಮ	ಬಾಣ	ದ್ರವ	ಮಣಿ
14n	ಮುದ್ರೆ	ದುಡ್ಡು	ಹತ್ತು	ಗಾತ್ರ	Hಟ	ಓಡು	ಬಿಲ
15o	ಮತ್ತು	ವೈದ್ಯ	ಬೆಕ್ಕು	ರವಿ	ಆಮೆ	ಗಂಡ	ಹುಳು
16p	ಮಣ್ಣು	ಮದ	ತೋಟ	ಮೂರು	ಮಣೆ	ಮರೆ	ಮೂಢ
17q	ಅಮ್ಮ	ಒಂದು	ಹುಲಿ	ದಾರಿ	ಹಕ್ಕು	ವಾದ	ಅಳು
18r	ಡಬ್ಬ	ಪ್ರಜ್ಞೆ	ದರ್ಪ	ಗಾಢ	ಕಾಲು	ಧೈರ್ಯ	ಗಂಡು
19s	ಗೊಂಬೆ	ಹಿಮ	ಅಟ್ಟ	ದಿಂಬು	ಅನ್ನ	ಪೆಟ್ಟು	ಬುದ್ಧಿ
20t	ಬೇವು	ಅಪ್ಪ	ರಾತ್ರಿ	ರೆಪ್ಪೆ	ದಾರ	ಮುಖ್ಯ	ಮಂತ್ರ
21u	ಮೊತ್ತ	ಕಟ್ಟು	ಮೊನ್ನೆ	ವಾಣಿ	ಕರ್ಮ	ಎಂಟು	ರಣ
22v	ಪ್ರಾಣ	ದೊಣ್ಣೆ	ಕೊಂಬೆ	ಅಡ್ಡ	ದಾಹ	ªÁåಪ್ತಿ	ಒತ್ತು
23w	ಮಂಕು	ಪ್ರಭು	ಹಗ್ಗ	ಬೇರು	ಹಿಟ್ಟು	ಗಾಯ	ಹುಬ್ಬು
24x	ಕವಿ	ಯಮ	ಮೋಹ	ಪತ್ರ	ವರ್ಗ	ನೋಟ	ನುಲಿ
25y	ಇದು	ಅಂತ್ಯ	ಹೆಣ್ಣು	ರಂಗ	ಮಾವ	ರಾಗ	ಲಡ್ಡು

HINDI LOW FREQUENCY WORD LISTS

	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8	List 9	List 10
1	बहू	अलाप	बैरा	बदल	आगा	अंजाम	आत्मा	अगर	बंधा	आला
2	बंजर	अटक	बंधक	बाधा	बंधु	बाडा	अंदाज	अनार	भवन	अधर
3	बारह	बहा	बंदर	बौना	बेदाग	बहाव	बयान	बाहर	डाबर	अंदर
4	बयाँ	बॅलन	बवाल	भरत	भुवन	बंधन	बॅटा	बाजू	दौरा	अटल
5	भाडा	भाग्य	भालू	दायर	दमन	बेताब	बुआ	बाला	गौरव	बाबुल
6	दाता	बौंरा	भजन	डंडा	दातुन	दीला	दावा	बामन	जरा	बना
7	दादी	बूरा	भुजा	धीमा	दवा	धीरज	धावा	बढाव	जंगल	भाला
8	देखा	दादी	डाका	गदर	धंधा	गवाह	डॉरी	भला	मैदा	भानु
9	ढोलक	डाला	दामाद	गोकुल	धारा	गुलाल	गला	बूढा	मकर	भाग्य
10	घुमाव	दीपक	धमाल	गुण	गाजर	जटा	गुरू	दमा	मरण	भाई
11	जहाज	गेहूँ	जवान	जंतु	घातक	जतन	जवान	धारण	मौका	भोंपू
12	लावा	घूँघट	लागत	लगाम	गोदाम	लालन	जमाव	गुना	नरम	बॉली
13	लाया	जागो	माता	लिखा	गोता	लहर	महा	जिला	पगार	दवात
14	मंजन	जोडा	मावा	मेवा	जंतर	लेखक	नाला	मादक	पैंदा	दोना
15	नदी	लेखन	नयाब	नवल	जवाब	लोहार	नादान	मैय्या	पलंग	गंगा
16	नारद	मंजा	नवाब	पाना	लगान	मजार	नकद	मरू	पंजा	लाली
17	नेता	मेला	निखिल	पकड	मंदा	नाता	पराग	पीपल	प्रकार	लगाव
18	निधि	मुराद	नियम	पलट	मात्रा	पालन	रहन	पौधा	राहू	मजाक
19	निपुण	मुर्गा	पागल	रानी	मिलाप	पांया	राणा	प्रबंध	राजा	मुकर
20	पहाड	पर्वत	पावन	तंबू	पडाव	पीपा	रेणु	तांगा	तालू	परे
21	पहल	पूरण	राहत	त्रिगुण	परख	पूजा	ताना	तमाम	तरुण	पवन
22	पोखर	प्रहार	रेखा	उपज	पुकार	पुर्जा	तोता	ताऊ	तीखी	तलब
23	प्रदान	तडप	रेला	वजन	तपन	ताई	वॅतन	वार्ता	तीतर	तनु
24	रुकाव	तरुण	तालाब	युवा	उधार	वहन	विधि	वधू	यात्रा	तीखा
25	ताजा	तवाह	विवाद	युवक	वादा	यात्री	यदा	विदित	युगल	तुला