THE EFFECT OF FILTERED SPEECH ON SPEECH IDENTIFICATION SCORES OF YOUNG NORMAL HEARING ADULTS

***Vijay Kumar Yadav Avilala, **Prashanth Prabhu P., & ***Animesh Barman**

Abstract

Filtered speech material would help to understand the importance of different spectral energy in perception of speech. Spectral energy of different speech sounds varies due to variation of the fundamental frequency (F0), first formant (F1) and second formant (F2) values. These are essential cues to understand speech. The F1 and F2 values are determined by the tongue height and tongue advancement respectively. These formants can alter due to the change in volume or size of the oral structures and also the language used. Hence, the present study aimed at assessing the speech identification scores for four low-pass (800, 1200, 1500 & 1700 Hz) and four high-pass (1700, 2100, 2500 & 3000 Hz) cut off frequencies filtered speech in 30 young normal hearing adults in the age range of 17- 25 years. The spectral modifications were done using Adobe Audition Software for the phonemically balanced words in Kannada developed by Yathiraj and Vijayalakshmi (2005). The speech identification scores were determined at all the eight cut off frequencies. Results revealed that there is lowering of the cut off frequencies at which 70% speech identification scores are obtained for low-pass filtered stimuli in Kannada compared to English (Bornstein, Wilson & Cambron, 1994). This difference could be because of the low F1 values and higher low frequency spectral energy in Kannada spoken language compared to English language. There is no difference in high-pass cut off frequency filtered speech. The results suggested that the spectral information between 1200 Hz and 2100 Hz is more important in perception of speech in Kannada language and also highlight the selection of appropriate cut off frequency for filtered speech for clinical use.

Key Words: High-pass cut off frequency, Low-pass cut off frequency, Spectral Modifications.

Speech is a complex signal which encompasses spectral information in the frequency range of 300 to 3000 Hz. Speech consists of different classes of speech sounds like vowels and consonants. Spectral energy of vowels is more concentrated at low to mid-frequencies, whereas consonants have spectral energy at wide range frequencies. Within the consonants, nasal speech sounds are more of low frequencies, plosives usually have more spectral energy at mid frequencies and fricatives usually have high frequency spectral energy concentration.

Due to the spectral variations of speech sounds, speech materials have become an indispensable tool in clinical evaluation. They could be used to determine the extent to which a person has disruption in the perception of complex signals like speech (Wilson & Margolis, 1983). Speech materials are also being used in selection and prescription of amplification devices and in rehabilitation (Risberg & Martoni, 1972). Beattie, Edgerton and Svihovec (1977) compared the slopes of performance intensity function of NU No. 6 and CID W-22 speech materials and reported that the slope was 4.2% and 4.6% respectively. It suggests that the different test materials can also yield different performance-intensity functions.

*Audiologist Grade I, All India Institute of Speech and Hearing,(AIISH), Mysore- 06, e-mail: avk.aud@gmail.com, **Audiologist Grade I, AIISH, Mysore- 06, e-mail : prashanth.audio@gmail.com, &***Reader, Department of Audiology, AIISH, Mysore- 06, e-mail: nishiprerna@rediffmail.com.

Speech contains both spectral and temporal information that is important for perception of speech. However, there are variations across languages in the way these spectral/temporal cues contribute in perception of speech. There are spectral variations like differences in formant frequencies and temporal variations like changes in speaking rate across languages. However, it might be complicated to study the influence of this information in speech identification scores by varying both the parameters. Hence, a systematic study is required to see the spectral or temporal influence on speech identification scores by keeping one parameter unaltered. Thus, by varying the spectral properties of speech (like use of filtered speech material) one can determine the contribution of different spectral energy in perception of speech in different languages.

Spectrally modified speech stimuli like filtered speech have been used as a monaural low redundancy test to assess Auditory processing disorder (Bocca, Calearo & Cassinari, 1954). Filtered speech helps to understand the contribution of different frequencies in perception of speech (Bornstein, Wilson & Cambron, 1994). The degradation of the speech signal that is produced by filtering is affected by the cut-off frequency of the filter and also by the rejection rate of the filter. Bornstein et al. (1994) reported that, (1) for low-pass filtered speech materials, the word recognition performance decreased with the lowering of the cut off frequencies; (2) for high-pass filtered speech materials, the performance decreased as the high pass cut-off frequency increased, and (3) for both low-pass and high-pass filtered speech material, the word recognition performance decreased as the steepness of the rejection rate of the filter increased.

Bornstein et al. (1994) observed that individuals with normal hearing could obtain 70% correct scores with low-pass cut-off frequencies of 1500 Hz and high pass cut off frequency of 2100 Hz. These results suggest that frequencies between 1500 Hz and 2100 Hz are more important in perception of speech. The spectral information above 2100 Hz and below 1500 Hz though important but may not provide adequate information for the perception of speech. The effect of cut-off frequency mainly depends on the spectral energy of the speech sounds and it might change with the language used.

The variations in spectral energy might alter due to the structural variations of the oral cavity. These could lead to the variations in first formant (F1) and second formant (F2), as F1 is determined by the frontal volume of the oral structure and F2 is by the volume of the back cavity. Savithri, Jayaram, Venugopal and Rajasudhakar (2007) have reported the mean F1 (497 Hz) and F2 (1596 Hz) values for vowels in Kannada language. Peterson and Barney (1952) have observed the mean F1 (538 Hz) and F2 (1560 Hz) values for vowels in English language. The comparison of F1 and F2 values reported in both the studies shows that the mean F1 of vowels in Kannada are lower in frequencies than the English language. The mean F2 of vowels is higher in Kannada compared to English language. Sairam (2002) reported that the Indian languages like Kannada have more energy in the low frequency region, compared to the western spectrum. Most of the studies reported in the literature regarding spectral variations of speech on speech perception (filtered speech) have used English language and those results cannot be adapted directly to other languages, especially Indian languages like Kannada. There is a need to develop data base for filtered speech stimuli in Indian languages as well. Hence, the current study aimed to observe how different spectrally varied speech stimuli can affect the speech identification scores.

Objectives of the Study: The main objective of this study was to determine the effect of spectrally modified speech (filtered speech) on the speech identification scores in normal hearing adults. The second objective was to find out the cut-off frequencies at which young normal hearing adults can obtain 70% speech identification scores for both low- and high-pass cut off frequency filtered speech.

Method

Participants: Thirty young adults (mean age 21.5 years) having normal hearing (hearing sensitivity less than or equal to 15 dB HL) participated in the study. The participants had no history of any otologic and neurologic symptoms. The participants had 100% speech identification scores (SIS) in quiet for monosyllables (Mayadevi, 1974) presented at most comfortable level. They had normal tympanogram with presence of acoustic reflexes at normal levels. Speech perception in Noise (SPIN) test was administered to rule out Auditory Processing Disorder and all the participants had speech identification scores greater than 60% at 0 dB SNR (Orchik & Burgess, 1977).

Stimuli: In the present study, phonemically balanced word lists developed by Yathiraj and Vijayalakshmi (2005) in Kannada were used for spectral modification to determine speech identification scores in young adults having normal hearing. The Yathiraj and Vijayalakshmi (2005) test material in Kannada has a total of four lists having 25 words in each list. These materials were standardized on young normal hearing adults. Each list was filtered using low-pass cut-off frequencies of 800, 1200, 1500 and 1700 Hz; and high-pass cut-off frequencies of 1700, 2100, 2500 and 3000 Hz using adobe audition software and the attenuation rate of 115 dB/octave using Butterworth filters. Each unmodified word was selected and filtered separately for each cut-off frequency. These four lists were further randomized using random table to make eight lists, to avoid any practice effect.

The words lists were recorded by a native Kannada female speaker. The recording was done on a Pentium Dual Core laptop using a unidirectional microphone kept at a distance of 10 cm from the speaker's mouth for recording. The recording was done using Adobe audition software (Version 2) using a 32-bit analog-to-digital converter at a sampling rate of 44.1 kHz. The recorded signal was normalized so that all the words had the same intensity. A calibration tone of 1 kHz was recorded prior to the list.

Procedure: To estimate the pure tone thresholds and speech identification scores, a calibrated dual channel diagnostic audiometer GSI-61 with TDH-39 headphones housed in MX-41/AR ear cushions was used. A Radio ear B-71 bone vibrator was used to estimate the bone conduction thresholds. Pure tone testing was done using Modified Hughson and Westlake procedure (Carhart & Jerger, 1959). Speech identification scores of these participants were determined using the monosyllables developed by Mayadevi (1974). Speech identification scores were obtained at most comfortable level. Immitance evaluation (tympanometry and acoustic reflex testing) was carried out using 226 Hz probe tone. A calibrated Middle ear Analyzer (GSI-Tympstar V 2.0) was used for the same. Acoustic reflexes were checked at 500, 1000, 2000 and 4000 Hz tone for both ipsilateral and contralateral. Speech Perception in Noise test was administered at 0 dB SNR using monosyllables developed by Mayadevi (1974). The speech material recorded for the study was played using Adobe Audition (Version 2.0) software. The signal was routed through a Core 2 Duo Computer to the tape and auxiliary input of a clinical audiometer through TDH-39 headphones with MX-41/AR cushions. The intensity of the presentation level was controlled from the audiometer.

The individuals were instructed to give a written response and the speech identification scores were determined for each cut-off frequency. The participants were also informed that they could guess the test items in case they were not very clear. Half of the participants were tested in right ear first and for the remaining participants were tested in left ear to avoid ear effect. The presentation of the stimuli was randomized for all the eight lists to avoid practice and order effect. The number of corrected responses was calculated. The speech identification scores were calculated using the formula given below:

$$SIS = \frac{\text{Optained number of responses}}{\text{Total number of responses}} * 100$$

Results

The speech identification scores obtained for different filtered speech was noted. The mean and standard deviation of the speech identification scores were calculated. The results obtained at different cut-off frequencies for 30 individuals (60 ears) with normal hearing are depicted in Table 1. It is evident from the table that there is an increase in speech identification scores with increase in low-pass cut-off frequency and with decrease in high-pass cut off frequency. The participants obtained greater than 70% scores for low-pass cut-off frequency of 1200 Hz or higher and high-pass cut off frequency of 2100 Hz or lower. The scores gradually decreased from 1700 to 800 Hz low-pass cut off frequency. However, there is a gradual decrease in speech identification scores from 1700 to 2500 Hz high-pass cut off frequency and then the scores sharply deteriorated for 3000 Hz high pass cut off frequency.

Cut-off frequency	n	Mean	SD	Range (Min - Max)
800 Low-pass	60	65.9	5.64	52-72
1200 Low-pass	60	75.86	3.71	72-88
1500 Low-pass	60	84.9	3.26	80-92
1700 Low-pass	60	90.9	2.76	88-96
1700 High pass	60	90.4	3.15	80-92
2100 High pass	60	77.86	4.42	72-88
2500 High pass	60	65.73	7.02	40-76
3000 High pass	60	36.4	4.49	24-48

Table 1: *Mean, standard deviation and range of SIS obtained for filtered speech stimuli at different cut off frequencies obtained from 60 ears of 30 individuals with normal hearing.*

A repeated measures ANOVA was carried out to test whether the speech identification scores were significantly different across different low and high cut-off frequencies. The results of repeated measures ANOVA shows that there was significant difference in speech identification scores across low cut off frequencies [F (3, 177) = 395.83, $p < 0.01$] and high cut off frequencies [F (3,177) = 924.67, $p < 0.01$].

A test of Bonferroni's multiple pair-wise comparison was carried out to determine between which two cut off frequencies the speech identification scores differed significantly from each other. The results of the Bonferroni's test revealed that the scores of all the cut-off frequencies for both low-pass and high-pass words differed significantly from each other at level of significance of $p < 0.01$ as depicted in Table 2 and 3.

identification scores for 1500 Hz low-pass and 2100 Hz high-pass cut-off frequencies. In the present study, it was found that the speech identification scores reached 70% at 1200Hz low-pass cut-off frequency itself and at 2100 Hz for high-pass cut-off frequency. This difference is mainly due to the difference in languages used. The discrepancy in the low-pass cut-off frequency for Kannada (1200 Hz) in comparison with English (1500 Hz) could be due to the predominance of low frequency information in Kannada language. Sairam (2002) reported that the spectral energy concentration of speech sounds in Kannada is more at the low frequency region compared to Western speech spectrum. This was attributed to the frequent occurrence of long vowels in Kannada which carries

Cut off frequency	1200 Hz low-pass	1500 Hz low-pass	1700 Hz low-pass
800 Hz low-pass	Significant (p < 0.01)	Significant (p < 0.01)	Significant (p < 0.01)
1200 Hz low-pass	-	Significant (p < 0.01)	Significant (p < 0.01)
1500 Hz low-pass	-	-	Significant (p < 0.01)

Table 2: *Bonferroni's multiple comparison of filtered speech for low-pass cut off frequencies.*

Cut off frequency	2100 high-pass	2500 Hz high-pass	3000 Hz high-pass
1700 Hz high-pass	Significant (p < 0.01)	Significant (p < 0.01)	Significant (p < 0.01)
2100 Hz high pass	-	Significant (p < 0.01)	Significant (p < 0.01)
2500 Hz high-pass	-	-	Significant (p < 0.01)

Table 3: Bonferroni's multiple comparison of filtered speech for high-pass cut off frequencies.

Discussion

The results of the present study show that the speech identification scores improved with increase in low-pass cut off frequencies and with decrease in high-pass cut off frequencies. The subjects obtained speech identification scores of 70% at low-pass cut off frequency of 1200 Hz and at high-pass cut off of 2100 Hz.

Bornstein et al. (1994) reported that individuals with normal hearing obtained 70% speech

more of low frequency information. The decreased low-pass cut off frequency in Kannada could also be because of lowered F1 values in Kannada compared to English which is important for consonant perception as observed by Savithri et al. (2007) in Kannada and Peterson and Barney (1952) observed in English language. The variation in F1 values could be because of changes in tongue position and volume of the back cavity (Savithri et al., 2007).

The present study shows that there was no difference in high- pass cut off frequency between Kannada and English reported by Bornstein et al. (1994). The mean F2 values for vowels in Kannada (Savithri et al., 2007) are higher than the F2 values for English reported by Peterson and Barney (1952). The mean F2 values for both Kannada and English were less than 1700 Hz. The high-pass cut off frequencies used in the present study is greater than 1700 Hz. Thus, the variation in F2 between the languages may not have affected the high-pass cut off frequencies. The result of the present study suggests that spectral energy information greater than 1200 Hz and lesser than 2100 Hz along with their auditory closure abilities used to obtain 70% speech identification scores. This suggests that the energy concentration between 1200 Hz and 2100 Hz are essential to perceive speech in Kannada language.

Conclusions

The present study made an attempt to find out the effect of spectral variations on perception of speech in young normal hearing adults. The study showed that the spectral information between 1200 Hz and 2100 Hz are important for perception of speech in Kannada. It was also found that slightly lower low cut off frequency is important to perceive the speech in Kannada compared to English language.

Implications of the study

This data can be used to carry out further studies in perception of filtered speech. It highlights the importance of studying filtered speech in different languages as the present study shows difference in lower cut-off frequencies compared to English languages. It also suggests the lowering of cut off frequency in filtered speech of Kannada language might be used to assess auditory processing disorders.

References

Beattie, R. C., Edgerton, B. J., & Svihovec, D. V. (1977). A comparison of the Auditec of St. Louis cassette recordings of NU-6 and CID W-22 on a normal hearing population. *Journal of Speech and Hearing Disorders, 42,* 60-64.

Bocca, E., Calearo, C., & Cassinari, V. (1954). A new method for testing hearing in temporal lobe tumors. *Acta Otolaryngologica, 44,* 219-22.

Bornstien, S. P., Wilson, R. H., & Cambron, N. K. (1994). Low- and high-pass filtered Northwestern University Auditory Test No. 6 for monaural and binaural Evaluation. *Journal of American Academy of Audiology, 5,* 259-264.

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

Mayadevi. (1974). *The development and standardization of a common speech discrimination test for Indians.* Unpublished dissertation submitted to University of Mysore, in part fulfillment of Masters Degree in Speech and Hearing.

Orchik, D. J., & Burgess, J. (1977). Synthetic sentence identification as a function of the age of the listener. *Journal of Speech and Hearing Research, 20,* 322-336.

Peterson, G. E., & Barney, H. E. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America, 24,* 175-184.

Risberg, A., & Martony, J. (1972). A method for the classification of audiograms, In Fant, G (Ed), *Speech Communication Abilities and Profound Deafness.* Washington, DC, AG Bell Association, 135-139.

Sairam, V. V. S. (2002). Long term average speech spectrum in Kannada. *Unpublished independent project submitted to University of Mysore, in part fulfillment of Masters Degree in Speech* and Hearing.

Savithri, S. R., Jayaram, M., Venugopal, M. B., & Rajasudhakar, R. (2007). Base of articulation of 13 Indian languages. *Journal of All India Institute of Speech and Hearing, 26,* 12-21.

Wilson, R. H., & Margolis, R. H. (1983). Measurements of auditory thresholds for speech stimuli. *Journal of Speech and Hearing Research, 11,* 256-270.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *Phonemically Balanced Word List in Kannada.* Unpublished departmental project in Department of Audiology, All India Institute of Speech and Hearing, Mysore.

Acknowledgements

The authors acknowledge with gratitude Dr. Vijayalakshmi Basavaraj, Director, All India Institute of Speech and Hearing, Mysore for permitting to conduct the study at the institute. We also like to thank Dr. Asha Yathiraj, Professor of Audiology, All India Institute of Speech and Hearing, Mysore, for permitting us to use the word list. The authors also like to acknowledge the participants for their cooperation. thank Dr. Asha Yathiraj, Professor of Audiology, All India Institute of Speech and Hearing, Mysore, for permitting us to use the word list. The authors also like to acknowledge the participants for their cooperation.